03043

4
2

GENERALIDADES DEL ANALISIS FACTORIAL

Programa de Especialización en Estadística Aplicada

Unidad Académica de los Ciclos Profesional y de Posgrado

Colegio de Ciencias y Humanidades

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

Tema de Especialidad presentada por :

CORINA MARGARITA CUEVAS RENAUD

Directora de la Tesina :

Dra. Guillermina Eslava Gómez

1996

AGRADECIMIENTOS

A la Dra. Guillermina Eslava Gómez por haber dirigido este trabajo con gran responsabilidad.

A los miembros del jurado:

Presidente: Dr. Serafín Mercado Domenech
Vocal: Dra. Guillermina Eslava Gómez
Secretario: Dr. Ignacio Mendéz Ramírez
Suplente: Dra. Silvia Ruíz Velasco Acosta
Suplente: Dr. José Rubén Hernández Cid

Por sus valiosos comentarios en la revisión de este trabajo.

A todos Uds. gracias por su calidad humana y su profesionalismo.

INDICE

CAPITULO 4

ROTACION FACTORIAL

CAPITULO 5

EJEMPLOS

CAPITULO 6

ANEXO A

INTRODUCCION

En este trabajo se abordan los conceptos fundamentales del análisis factorial desde una perspectiva matemática de fácil comprensión. Se pretende que sea de utilidad para quienes hacen uso del análisis factorial, principalmente en el campo de las Ciencias Sociales.

Desde principios del siglo XX con los trabajos desarrollados por Spearman en el año de 1904 se inició el estudio de una técnica estadística, conocida actualmente como Análisis Factorial, a los pocos años de que la Psicología se constituyó como Ciencia (hacia fines del siglo XIX). Podemos considerar que los trabajos de Spearman fueron encaminados a darle a la Psicología el caríz de una ciencia cuantitativa. La Psicología estudia la conducta del ser humano, entendiendo por conducta el conjunto de respuestas con las que éste se comporta ante los estímulos que provienen tanto de su interior como del exterior, determinando sus condiciones y leyes. La conducta del ser humano está determinada por un gran número de factores o aspectos que no pueden medirse "directamente", como la inteligencia, la personalidad, las actitudes, el consumo de fármacos, el desarrollo psicomotor de un niño, el síndrome de las victimas de abuso sexual, síntomas psicóticos, intereses vocacionales, entre otros. Cada uno de estos constructos está integrado por un grupo de variables indicadoras, manifiestas u observadas, que se relacionan entre sí para dar origen a las estructuras conceptuales o variables latentes.

En una sociedad cada vez más compleja, la Psicología ha ido ocupando lugares importantes en la prevención y solución de problemas fundamentalmente sociales. Por lo cual la búsqueda de los conocimientos y las técnicas que le permitan abordar adecuadamente estos problemas es imprescindible. Particularmente el análisis factorial constituye una de las técnicas estadísticas multivariadas de mayor aplicación en el campo profesional y de investigación del psicólogo, actividades tales como la construcción y validación de instrumentos y la construcción de modelos principalmente, demandan la presencia del análisis factorial.

El extraordinario avance en los últimos años en materia de informática ha permito un gran desarrollo de paquetería (software) en Estadística, lo que ha ocasionado entre otras cosas el uso indiscriminado del análisis factorial y de otras muchas técnicas estadísticas de quienes desconocen los supuestos matemáticos que subyacen a dichas técnicas, llegando a considerar que el análisis de componentes principales es lo mismo que el análisis factorial, así como una serie de dogmatismos tales como que sólo los factores con un eigenvalor mayor a uno son válidos, el hacer estimaciones en análisis factorial con componentes principales, llevar a cabo mediciones de variables observadas, que en número es mayor al número de sujetos en quienes se realizan las observaciones o mediciones y tomar solamente en cuenta que el nivel de medición de las variables sea al menos intervalar como único supuesto del análisis factorial, entre otros.

En el capítulo 1 se presentan los antecedentes históricos del análisis factorial, su desarrollo a través del tiempo, presentando las contribuciones más importantes al mismo. Se dedica una buena parte en este primer capítulo al trabajo desarrollado por Spearman replicándose su estudio del "*factor general*" de la inteligencia así como los enfoques de algunos de los principales exponentes del análisis factorial como Thurstone, Cattell, Bartholomew y Jöreskog. En el capítulo

2 se presenta el concepto de análisis factorial, su modelo, sus principios, sus elementos básicos, tales como variables manifiestas, variables latentes o factores, cargas factoriales, eigenvalores, comunalidades, varianza única o residuos, entre otros, así como sus principales diferencias con el análisis de componentes principales. En el capítulo 3 se presentan cuatro de los métodos de estimación más empleados:
1) mínimos cuadrados, 2) mínimos cuadrados generalizados, 3) mínimos cuadrados ponderados y
4) máxima verosimilitud, haciendo mayor énfasis en este último método.

En el capítulo 4 se presenta el concepto de rotación factorial: ortogonales y oblicuas, se reproduce un ejemplo desarrollado por Bhartolomew en donde se aprecia cómo, para ese caso, la rotación oblicua recupera más fielmente la información de la matriz de correlación original para un grupo de datos hipotéticos. En este mismo capítulo se incluyen los temas de selección del modelo y de los puntajes factoriales (*factor scores*). En el capítulo 5 se hace una aplicación del análisis factorial haciendo uso de tres paquetes estadísticos: LISREL, SPSS y SAS. Finalmente en el capítulo 6 se presentan las conclusiones generales del trabajo. Se incluye un anexo, con los datos del ejemplo de la investigación sobre inteligencia de Spearman y del ejemplo hipotético de Bartholomew.

CAPITULO 1

ANTECEDENTES HISTORICOS.

En este capítulo se presenta una semblanza histórica del surgimiento y desarrollo del análisis factorial (Mulaik, Stanley 1986), iniciándose con Spearman en 1904 a quien se le considera el padre del análisis factorial, y hasta las últimas aportaciones de la década de los años 80'.

Charles Spearman al estudiar la medición de la inteligencia empleó el análisis factorial bajo el supuesto hipotético de que si dos habilidades están correlacionadas entre sí, en cierta medida, cada una de ellas incluye dos factores: un factor común a ambas (que determina la correlación entre esas habilidades) y un factor específico a cada una (que determina la diferenciación entre esas habilidades). Para ello se abocó al estudio de la relación entre *inteligencia general* (General Intelligence) y *discriminación sensorial* (Sensory Discrimination), la que llamó *Psicologia Correlacional* dentro del área experimental de la Psicología. Esto lo realizó con objeto de proporcionar información que permitiera la definición objetiva de la inteligencia y sobre todo encontrar el procedimiento de su medición (Spearman, 1904).

Para ello llevó a cabo 5 series de experimentos en poblaciones de adolescentes, considerando en su estudio 4 áreas de discriminación sensorial, siendo dichas áreas las siguientes:

1) *Pitch* (discriminación de tonos).
2) *Sight* (discriminación visual).
3) *Weight* (discriminación de pesos).
4) *Intelligence* (efecto de la enseñanza). Desafortunadamente no surgió este efecto en las mediciones, excepto las conclusiones que arrojaron las otras áreas en los propios experimentos.

Demostró como la edad y el sexo eran irrelevantes para su estudio, presentando una amplia justificación al respecto en este trabajo.

El experimento serie I se llevó a cabo con 24 adolescentes, sin ninguna selección y se midió en ellos discriminación visual, de pesos y de sonido. Cada adolescente fue evaluado en estas áreas de forma separada en la propia casa de Spearman en los días de descanso que ellos tenían .

El experimento serie II se efectuó en el mismo escenario que el anterior, pero ahora fue con 36 adolescentes de los restantes. Este grupo de adolescentes fueron evaluados sólo en discriminación de sonidos y la evaluación fue colectiva.

El experimento serie III se efectuó solamente en discriminación visual y de pesos teniendo como escenario la propia escuela participando sólo 24 adolescentes de un total de 37. La aplicación en este experimento no fue colectiva, sino individual.

El experimento serie IV se efectuó solamente en discriminación de sonidos, el escenario fue en la misma escuela preparatoria que para el experimento anterior, pero en condiciones experimentales más propicias que las anteriores. La capacidad de discriminar un tono musical es una capacidad sensorial suceptible de ser evaluada de manera colectiva, y así lo hizo con 33 adolescentes.

El experimento V se llevó a cabo con un total de 26 hombres y mujeres adultos. Con estos sumados a los anteriores y considerando la exclusión de algunos casos y/o el retiro de los mismos dió un total de 123 sujetos. En este experimento la evaluación fue individual exactamente igual al descrito en el experimento serie I.

El estudio de la inteligencia señala Spearman es una materia díficil y en sus estudios sobre la misma no hizo ninguna suposición previa de qué clase de actividad mental podría estar subyacente a la variable latente llamada inteligencia; de ahí que este estudio se considere exploratorio. Parte de la evidencia de la existencia de cuatro diferentes clases de inteligencia en función de varias habilidades:
1) en función del lugar que ocupa el alumno en su salón de clase (del mejor al peor) estando claramente representada por lo que el llamo "*Present Efficiency*", en materias del área denominada Clásicas (Classics) que incluyeron Latín y Griego así como de Matemáticas. Particularmente en los experimentos series III y IV se estudió este tipo de inteligencia.
2) la que él llamó "*native capacity*" que bien podría traducirse como capacidad inata que nada tiene que ver ni con la edad ni con el sexo.
3) Representada y medida por las impresiones generales producidas sobre otras personas, que se traducen por ejemplo en la clasificación que los maestros hacen de sus alumnos en "*brillantes*", "*regulares*" y "*malos*".
4) Inteligencia estimada (no medida directamente, sino por medio de una prueba) se conoce como "*common sense*", es decir, sentido común. En el experimento serie I se trabajó con este tipo de inteligencia.

Spearman señala que de los 5 experimentos que realizó, el experimento serie IV fue el mejor, el más preciso. Por ello quizás es el que reportan la mayoría de los libros tanto de Psicología como de Análisis Factorial y justamente es el que presentamos a continuación.

Como recordaremos, sólo se midió la capacidad de discriminación de sonidos y se correlacionó con la segunda clase de inteligencia, la capacidad inata medida a traves de las siguientes variables manifiestas: *Classics, French, English y Mathematics*. Obteniendo la siguiente tabla de correlaciones Kendall (1980, p. 47):

MATRIZ DE CORRELACION

variables manifiestas	1	2	3	4	5	6
1 Classics	-	.83	.78	.70	.66	.63
2 French	.83	-	.67	.67	.65	.57
3 English	.78	.67	-	.64	.54	.51
4 Mathematics	.70	.67	.64	-	.45	.51
5 Discrimination	.66	.65	.54	.45	-	.40
6 Music	.63	.58	.51	.51	.40	-

Tabla No. 4

Vale la pena mencionar que estas correlaciones no se obtuvieron con el estadístico de correlación producto-momento de Pearson, Spearman modificó la fórmula de dicho estadístico justificándose en el sentido de que con ello se eliminan los errores en las observaciones, obteniendo dos fórmulas de correlación a las cuales llamó teórica y empírica. Spearman observó en esta tabla de correlación como había un comportamiento aproximadamente proporcional de decrecimiento en las correlaciones de las variables manifiestas 1, 2, 3, 4, 5 y 6 justamente en este orden. Situación que llamó poderosamente su atención y de esta inquietud se iniciaron los primeros trabajos encaminados a explicar este comportamiento, surgiendo así el análisis factorial.

Con los datos de la matriz de correlación anterior, se procedió a procesarlos con el Paquete Estadístico para las Ciencias Sociales (SPSS) en su versión 5.0 para Windows empleando la subrutina "*Factor Analysis*", así como el método de extracción Componentes Principales (PC) sin rotar y extrayendo los factores que por omisión dá el SPSS (con eigenvalores mayores a uno, según el criterio dado por Kaiser-Guttman), obteniendo los siguientes resultados (en el anexo A se presentan las instrucciones dadas).

MATRIZ DE CARGAS FACTORIALES

VARIABLES	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5	FACTOR 6
X1	.93614	-.02740	-.02131	-.07756	-.13822	-.31196
X2	.89542	-.08294	.01878	.13384	-.38068	.16780
X3	.84162	-.00494	-.25115	-.45425	.09809	.11230
X4	.80379	.20086	-.39972	.34238	.19108	.00804
X5	.74220	-.56486	.26977	.09594	.21818	.02221
X6	.72286	.50060	.46424	-.01568	.09986	.03366

Tabla No. 5

Con los eigenvalores de 4.11, .62, .51, .36, .27 y .14 para los factores 1, 2, 3, 4, 5 y 6, así como un porcentaje de varianza explicada de 68.4%, 10.3%, 8.5%, 5.9%, 4.5% y 2.3% para los mismos factores respectivamente.

ESTADISTICAS FINALES

VARIABLE	COMUNALIDAD	FACTOR	EIGENVALOR	% de VAR.
CLASSICS	.87635	1	4.10592	68.4
FRENCH	.80178			
ENGLISH	.70832			
MATHEMATICS	.64608			
DISCRIMINATION	.55085			
MUSIC	.52253			

Tabla No. 6

De estos resultados se obtiene un sólo Factor llamado por Spearman *"Factor General"*, en el cual todas las variables tienen cargas factoriales >.72 el cual explica el 68.4% de la varianza total.

A continuación presentamos un breve resumen de los hallazgos obtenidos por Spearman en sus investigaciones a lo largo de 23 años (1904-1927).

La teoría de Spearman de la *"Unidad Universal de la función intelectual"*, como él la llamó en su primer hallazgo en 1904 es ahora modestamente llamada la *"Teoría de la Inteligencia de dos factores"* (*Two-factor theory of intelligence*). Un factor es el elemento común, *general*, y el otro es el factor peculiar para una tarea dada, *special*). La llave de la evidencia de Spearman fue que cada tarea mostró considerable consistencia, en sus correlaciones con otras tareas o ejecuciones. No igualdad, pero sí consistencia. Es decir, *clásicas* correlacionó más fuertemente con *matemáticas* (.70) que con *discriminación sensorial* (.66), pero *matemáticas* correlacionó más altamente con otras tareas o ejecuciones que con *discriminación sensorial*, por ejemplo con *francés* (.67) con *inglés* (.64) y con *música* (.51). *Discriminación sensorial* fue consistentemente una ejecución que correlacionó bajo, aunque de todas estas correlaciones sólo una fue alta la que correlacionó con *clásicas*, pero esto se debió a que *clásicas* fue la ejecución con la mayor correlación. Spearman mostró que sus datos fueron consistentes con la teoría de que todas las correlaciones estuvieron basadas en un sólo elemento común y algunas tareas o ejecuciones como *clásicas*, fueron altamente afectadas con la presencia de otros elementos, mientras que algunas ejecuciones, tales como *discriminación sensorial* no lo fue. Su argumento fue escencialmente un procedimiento matemático que simplifica una red de correlaciones.

Veamos ahora las aportaciones de otros estudiosos del análisis Factorial (Mulaik S. 1986). Raymond B. Cattell, estudió con Spearman. En sus investigaciones utilizó el método de diseño multivariado. Presentó la prueba gráfica (*Scree Test*) con respecto al número de factores a extraer.

Cyril Burt hacia fines de la segunda década de este siglo, encontró una forma más práctica para la solución de un sólo factor, mediante un conjunto de intercorrelaciones al cual llamó el método del *Centroide*. Este método básicamente consiste en encontrar una solución que convierta en un máximo *la suma de las cargas factoriales*. Posteriormente fue empleado y perfeccionado para más de un factor por L. Thurstone.

León Thurstone conjuntamente con Hull y Kelly trabajaron el análisis factorial múltiple o también llamado Análisis Multifactorial; pero de forma particular continuó trabajando en la medición de la inteligencia. En los esfuerzos que se hicieron para reformular a la Psicología como una ciencia cuantitativa iniciados por Spearman en 1904; sin duda alguna Thurstone junto con sus seguidores y alumnos ocuparon un lugar importante en el área de la Psicometría. Thrurstone fundó la revista *Psychometrika*, dominando por mucho tiempo su estilo de trabajo.

En una buena parte de la historia del modelaje estructural que se presentaba en al revista *Psychometrika*, se centró en el desarrollo de modelos estructurales conocido como *Análisis Factorial*. Sin embargo, no fueron los primeros en preocuparse por ello, los antecesores realmente fueron Spearman, Thomson, Garnett, Lederman y Burt.

Por los años 30' de este siglo, Thurstone hace su arrivo al Análisis Factorial constituyéndose en el centro de la actividad y desarrollo así como de sus usos y aplicaciones, iniciándose en las Islas Británicas y llegando a los Estados Unidos con gran fuerza hacia 1970.

El grupo de los "*Thurstonianos*" se dedicaron de manera indiscriminada a escribir sobre el Análisis Factorial como si el problema de la indeterminación del número de factores no existiera, lo cual fue motivo de preocupación para un grupo de analistas factoriales británicos.

Steiger y Schöneman tuvieron la impresión de que el grupo *Thurstoniano* negaba la existencia de este asunto, debido a que representaba un defecto fatal en su modelo factorial y consideraron que esta negación se debía a que la revista *Psychometrika* era la depositaria de todos los nuevos e importantes desarrollos del Análisis Multifactorial a partir de 1936 y debido a este vertiginoso desarrollo, ellos no estaban dedicados al estudio del problema de la no determinación y de un sólo factor general, que para cualquier duda recomendaban consultar la monografía de Wolfle llamada "*Factor Analysis to 1940*". En esta monografía, carente de las aportaciones originalmente hechas por sus iniciadores como Spearman y Garnet; Wolfle, dá una cuantas indicaciones sobre la no determinación tomadas de las discusiones que sobre esto hicieron Wilson, Spearman, Piaggio y Thomson.

En 1936, apareció publicado en *Psychometrika* este problema, por lo cual Thurstone y sus seguidores ya sabían lo que debían hacer. Sostuvieron que la forma de trabajar con el análisis factorial era el correcto, criticando los métodos de sus rivales como Spearman, Holzinger y Swinecizing y Hotelling. Además el grupo de los *Thurstonianos* estabán más preocupados por un buen número de problemas metodológicos y sus algoritmos de resolución tales como: ¿Cuál es la mejor manera de estimar las comunalidades?, ¿Cuál es la mejor manera de extraer factores?, ¿Cuál es la mejor manera de rotar factores, ¿Cómo se pueden obtener puntajes de factores comúnes?, ¿Cuáles son las propiedades invariantes reveladas por un análisis factorial? y ¿Cómo puede agilizarse el proceso de cálculo del análisis factorial?, entre otros.

Con respecto a los problemas metodológicos, Thurstone señaló que se podía hacer mal uso del análisis factorial insistiendo en que los procedimientos para su manejo eran los correctos. Fue sensible a la necesidad de que la Ciencia tiende a sobredeterminar los constructos hipotetizados de datos experimentales, enfatizando el hecho de que el número de factores debería ser mucho menor que el número de variables observadas o manifiestas. Thurstone igualmente aceptó el hecho de que él había usado el análisis factorial como exploratorio, realizando éste a "ciegas", lo cual pudo producir resultados carentes de interpretación, particularmente si numerosos factores fueron confundidos o desconocidos y se entremezclaron con las pruebas (variables manifiestas).

Sostuvo además, que el factor debe ser "*sobredeterminado*" con al menos 3 ó 4 pruebas. Dado que el número de pruebas es mayor en número que el de los factores obtenidos, se consideran factoriales simples, cayendo en grupos de 3 o más pruebas y cada grupo asociado a un factor diferente, por lo cual Thurstone creyó que el método de componentes principales o "*centroid factors*" no podría recobrar estos factores, pero si podría representar "*artifact*" matemáticos para posibles transformaciones. Con el argumento de que los factores al ser rotados tanto como fuera posible a una configuración simple en una disposición tal, permitirían identificar los factores correctos. De esta forma, Thurstone en la práctica, se acercó al análisis factorial confirmatorio, sólo que con su metodología lo podría haber hecho de manera burda.

Thurstone rechazó los modelos de Spearman, Holzinger y Hotelling, ya que todos ellos enfatizaban al análisis factorial como generador de factores generales. Stanley A. Mulaik (1986) opina a este respecto diciendo, que la razón de este rechazo fue porque Thurstone sospechaba que el factor general desarrollado por ellos en sus modelos contenían "*artifact*" matemáticos no muy bien definidos en el contexto científico, independientemente de la selección particular de las pruebas para las cuales sus modelos serían aplicados, existiendo con ello un divorcio entre la matemática y el contexto del problema. Así mismo, Thurstone sostuvo que componentes principales y ejes centroides, técnicas empleadas para la obtención de los factores generales en sus modelos, podían cambiar con facilidad si se incluían pruebas diferentes a las originales.

En el aniversario de los 50 años de la publicación de la revista *Psychometrika*, Mulaik Stanley (1986, *Psychometrika* Vol. 5 No. 1 pp. 23-33) escribe un artículo reportando la producción de gran número de avances metodológicos en el análisis factorial, de los cuales se presenta un resumen de las aportaciones más sobresalientes:

-Roff (1936, citado en Mulaik, 1986) en su artículo "*Some properties of the communality in multiple factor theory*", estableció que el cuadrado de la correlación múltiple de una prueba de predicción de p-1 variables es el límite menor de la comunalidad de la prueba.

-Eckart y Young (1936, citado en Mulaik, 1986) en su artículo "*The aproximation of one matrix by another of lower rank*" publicado en el tercer número de la revista *Psychometrika*, presentaron un teorema que lleva sus nombres en el cual señalan que para cualquier matriz real semipositiva A, dos matrices ortogonales U y V pueden ser encontradas de modo que L=U'AV es una matriz diagonal con elementos no negativos. Ha sido la mayor contribución a la teoría del álgebra matricial y fundamental para el análisis factorial. Esta es la llamada "*descomposición espectral*".

-Stephenson (1936, citado en Mulaik, 1986) con su artículo "*The foundations of Psychometry:Four factor systems*" introduce la técnica Q, la cual ha tenido cierta influencia posterior en algunas aproximaciones factoriales analíticas al análisis de conglomerados. La cual básicamente consiste en encontrar un método de clasificación de un gran número de sujetos en grupos diferentes.

-Engelhart (1936, citando en Mulaik, 1986) discutió la técnica del análisis de los coeficientes de senderos (*Path Analysis*).

-Lendermann (1936, citado en Mulaik, 1986) derivó su famosa desigualdad de Ledermann, la cual se refiere al menor rango con que una matriz de correlación puede ser aproximada, ajustando las comunalidades en la diagonal.

-Dwyer (1937, citado en Mulaik, 1986) discutió el análisis factorial en su artículo "*The determination of the loadings of a given test from the known factor loadings of other tests*".

FALTA PAGINA

No. 7

-Hendrickson y White (1964, citado en Mulaik, 1986) método PROMAX.

-Guttman (1953-1956 citado en Mulaik, 1986) escribió sobre la teoría de la imagen y problemas relacionados con las comunalidades. Mulaik (1986) señala que los artículos de Guttman contienen la más elegante álgebra lineal en la literatura del análisis factorial.

-Rao C.R. (1955, citado en Mulaik, 1986) desarrolló un algoritmo iterativo para estimar las comunalidades y los factores asociados, basado en la forma de considerar el análisis factorial en términos de correlación canónica.

-Harris (1962, citado en Mulaik, 1986) hace una síntesis de los métodos de Guttman y Rao C.R.

-Lord (1958, citado en Mulaik, 1986) estableció que el primero de los componentes principales de una batería de pruebas puede tener la máxima confiabilidad, en el mismo sentido que el coeficiente alpha de Cronbach; y así tener la mayor generalización para el campo de las pruebas de donde se extrajo como una muestra aleatoria de las pruebas, para cualquier composición o arreglo derivado de la batería de pruebas, esto se conoce con el nombre de confiabilidad.

-Kaiser y Caffrey (1965, citado en Mulaik, 1986) desarrollaron el análisis factorial alpha como un grupo de variables, el cual es considerado como un método de factorización inicial, donde las variables incluídas en el análisis son consideradas muestras de un universo de variables.

Aunque en 1960 ya existían un buen número de programas computacionales disponibles para realizar el Análisis Factorial, el problema de la estimación máximo verosímil no estaba técnicamente resuelto.

-Lawley (1940, citado en Mulaik, 1986) formuló ecuaciones para la estimación máximo verosímil de los parámetros del modelo factorial común.

-Un graduado del departamento de Estadística Matemática de Carolina del Norte, (1955, citado en Mulaik, 1986) desarrolló uno de los primeros algoritmos convergentes muy útil en el análisis factorial de máxima verosímilitud.

-Schönemann (1985, citado en Mulaik, 1986) trabajó desde 1965 con los datos obtenidos por el graduado anteriormente señalado, derivando las trazas y determinantes de las matrices que fueron relevantes para encontrar las soluciones máximo verosímiles, pero fue hasta 1985 que formalizó sus hallazgos dándolos a conocer a la comunidad científica.

Dos problemas fundamentales subsistían en el método de máxima verosimilitud: 1) Las ecuaciones para estimar la máxima verosimilitud de los parámetros del modelo de factor común eran obtenidas de la desviación de funciones matriciales, algunas de ellas involucrando determinates de matrices. Tratadas al nivel de álgebra escalar estas ecuaciones son muy complejas de trabajar y 2) las ecuaciones estimadas por el método de máxima verosimilitud no pueden ser resueltas algebraícamente, ya que requieren de soluciones numéricas iterativas. Este problema se solucionó gracias a los trabajos realizados por Karl Jöreskog. Su atención se centró en el problema de estimar los parámetros del modelo de factor común mediante la estimación de máxima verosimilitud. A partir de este momento, el énfasis se dió al análisis factorial confirmatorio muy a pesar de la inclinación del grupo de los Thurstonianos por el análisis factorial exploratorio.

La idea central de los métodos desarrollados por Jöreskog consistió en especificar los parámetros "*a priori*" de las hipótesis en el modelo del análisis factorial y estimar los parámetros así especificados. La idoneidad del modelo propuesto a las hipótesis se obtiene como resultado de estudiar las diferencias entre la matriz de covarianza original y la matriz de covarianza estimada. De esta forma los principales problemas de la estimación paramétrica en ambos análisis factoriales: exploratorio y confirmatorio fue esencialmente resuelto (Mulaik, 1986 op. cit. p. 28).

Sin embargo el problema de la no determinación (indeterminancia) persistía, hubo muchos intentos por resolverla figurando entre ellos los trabajos de Guttman, Shönemann, McDonald, nuevamente Shönemann y Steiger, Bock y Bergmann. Hasta que Golderberger y Jöreskog conjuntan estas aproximaciones unificando el análisis factorial, el análisis de la estructura de la covarianza y el modelaje de ecuaciones lineales estructurales en un sólo modelo general para resolver modelos de estructuras complejas vertido en el programa de computadora LISREL (*Linear Structural Relations*).

-Muthén (1984 citado en Mulaik, 1986) propone modelos factoriales mixtos, es decir, contínuos, dicotomizados y policotomizados derivados de la normal multivariada.

No obstante que el análisis factorial al paso del tiempo se consolida, surgen igualmente cuestionamientos, aún no resueltos, tales como los siguientes: ¿Puede ser considerado para su uso los métodos de análisis factorial confirmatorio para evaluar modelos causales, pero qué es causalidad, y qué se puede contestar a la pregunta sobre el efecto de la elección de los métodos para estudiar la causalidad?, ¿es posible estudiar causalidad en forma exploratoria?, ¿podemos siempre estudiar causalidad sin hacer suposiciones?, ¿podemos integrar probabilidad con determinismo en estudios de causalidad?.

Y tal como lo señala Mulaik (1986) para intentarlo es necesario combinar además de las matemáticas, la filosofía con la metodología.

ALGUNOS ENFOQUES DEL ANALISIS FACTORIAL

A continuación se exponen las ideas centrales de cuatro de los principales exponentes del análisis factorial: L. Thurstone, B. Raymond Cattell, David J. Bartholomew y K.G. Jöreskog y de una nueva "*filosofía*" sobre análisis de datos: J.P. Benzécri. Iniciamos esta exposición con el enfoque de Thurstone sobre el análisis factorial.

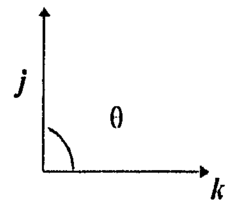
LEON L. THURSTONE

Thurstone conjuntamente con Hull y Kelly trabajaron el Análisis Factorial Múltiple o también llamado Análisis Mulifactorial de la Inteligencia; particularmente Thurstone continuó trabajando en la medición de la inteligencia. Thurstone señala que uno de los principales problemas en Psicología que ha llamado poderosamente la atención por muchos años es el problema de tratar de aislar o separar las habilidades mentales. Este es un problema desafiante en vista del hecho de que, en la naturaleza del caso, no podemos observar una habilidad mental aislada. Las observaciones que son aisladas en la Teoría Factorial son prácticamente siempre coeficientes de correlación o alguna medida comparable de asociación entre pares de ejecuciones. Se puede hablar acerca de las pruebas, pero la teoría factorial múltiple puede ser igualmente aplicada correctamente a una tabla de coeficientes de correlación que representan otra clase de variables. Se pueden mostrar algunos ejemplos ilustrativos; ya que pueden aplicarse no sólo a pruebas psicológicas sino además a variables tales como síntomas psicóticos, intereses vocacionales y pruebas de personalidad, entre otras.

El propio Thurstone, en su artículo "*La medición de la inteligencia*" (Thurstone 1947), señala que los métodos factoriales permiten determinar el número de factores linealmente independientes (que en el caso de la inteligencia se llaman capacidades) que han de postularse para explicar las correlaciones observadas en un determinado campo experimentalmente definido. El número de factores es el rango de la matriz de correlación. Cuando se han determinado las correlaciones de 30 tests, por ejemplo, y son de rango 5, el problema que resta es identificar estos 5 rasgos básicos responsables de las variaciones comunes en la batería de tests o pruebas. Aparece aquí una de las formas más interesantes de no determinación, cuya naturaleza puede ser ejemplificada geométricamente. En el modelo geométrico, que mostró ser muy útil en problemas de este tipo, cada test está representado por un vector. Estos vectores pueden simbolizarse en 3 dimensiones por largas agujas clavadas en un corcho central. El cuadrado de la longitud de cada aguja representa la fracción de la varianza que cada test comparte con los otros instrumentos de la bateria. Un vector muy corto, por lo tanto, indicaría que ese test tiene una parte reducida de su varianza propia dado que la comparte totalmente con uno o más de los otros tests. Si para explicar las correlaciones de la batería fueran suficientes 3 factores, toda la tabla de correlaciones podría ser simbolizada en un modelo físico como el que hemos descrito, en el cual cada aguja representa a un test. La correlación entre un par cualesquiera de tests estaría dada por el coseno del ángulo entre el par de vectores-tests. No se perdería, en la construcción del modelo, ninguna información y las correlaciones podrían, en su momento, ser reconstruídas a partir de los ángulos y longitudes de los vectores.

Igualmente señala que aquí, termina la cuestión matemática y se tiene un problema psicológico, ya que una vez comprobado experimentalmente que por ejemplo, el orden de las correlaciones es 3 se tienen que identificar, a partir del modelo, estas 3 capacidades fundamentales. El orden de las correlaciones es igual al número de dimensiones del modelo. En todos los problemas factoriales nos encontramos con la misma situación: las correlaciones determinan la configuración de los vectores sin ningún marco de referencia. Como se vé, el problema es doble, por un lado el geométrico y, por el otro, el de encontrar un marco teórico o de referencia adecuado y significativo en el modelo.

Geométricamente si los vectores se disponen en una configuración triangular, la ubicación más adecuada de los 3 ejes de referencia serán los 3 vértices del triángulo. Si no hay correlaciones entre estos 3 ejes fundamentales el modelo formará ángulos rectos, es decir, $r_{jk}=0$ cuando los vectores forman un ángulo de 90 °.



$r_{jk} = \cos(\theta)$

Figura No. 1

En caso contrario, sí se forman ángulos agudos u obtusos tendrán correlaciones positivas o negativas.

Cuando se soluciona el problema geométrico o algebraíco, se puede preparar lo que se denomina una matríz factorial que describe cada variable en función de los factores básicos. Por ejemplo Thurstone señala que en un test de opuestos o antónimos aparecerá con una gran saturación en uno de los factores verbales, y con saturación cercana a cero en razonamiento aritmético, por ejemplo. En otro caso, un test de razonamiento aritmético podrá estar saturado en varios factores, pero en fluídez verbal muy seguramente su saturación será cercana a cero.

Una vez que se han encontrado las capacidades básicas por medio de métodos analíticos, algebraícos o geométricos deberán someterse a una prueba final que Thurstone llama de "*Utilidad Científica*" la cual será dada por la significación de los factores o capacidades determinados. Si no hay una interpretación en este sentido, los instrumentos serán sólo estadísticos, de muy poca utilidad a la ciencia psicológica lo cual sería así en los campos de percepción, asociación y pensamiento.

Thurstone indica que otra forma de abordar el análisis anterior sería considerar primero, cuántos factores deben postularse para explicar las correlaciones observadas; dicho número será el orden de las mismas. Estos dos pasos cuidadosos y sucesivos deberán seguirse para posteriormente pasarse a lo que el propio autor señala como la "*prueba de fuego*", es decir, la de la "*valídez científica*" y, por ella debemos entender la capacidad de comprender el significado de cada factor como parámetro que determina el desempeño en el test, y de predecir la comprensión factorial de una nueva tarea mediante consideraciones puramente psicológicas sobre la naturaleza de los factores. Esto apenas nos acerca a la eventual explicación en el campo del conocimiento o de estudio.
Thurstone concluye en sus estudios sobre el Análisis Multifactorial de la Inteligencia que las capacidades mentales primarias identificadas con cierta "seguridad" son:

N : la capacidad numérica.
S : la capacidad visual, a la que se ha denominado factor espacial.
M : la capacidad de memorización (probablemente distinta de la memoria incidental y otros diversos tipos de memoria, como la visual).
V : la comprensión verbal.
W: la fluídez verbal.
I : la inducción.
P : la velocidad de percepción simple.

Así como un factor que favorece la rapídez del cierre perceptual, otro que favorece su flexibilidad, dos factores verbales que facilitan la asociación, uno de ellos limitado a la forma de la misma, y un factor de denominación que puede ser referido a la fluídez verbal. También parecen existir una o más capacidades intelectuales generales que subyacen a las capacidades mentales primarias y que explican, las correlaciones entre ellas, que según parece razonable, apoyarían el postulado sostenido por Spearman en sus primeros trabajos de un factor intelectual central sumado a las capacidades especiales.

El modelo propuesto por Thurstone es al igual al modelo planteado en la siguiente ecuación:

$$x = \mu + \Lambda y + e \qquad (1)$$

donde:

μ = constante que representa a la media.
x = vector de variables manifiestas u observadas.
Λ = matriz de cargas factoriales.
y = vector de variables latentes, también llamadas factores.
e = vector de variables específicas (residuos) de elementos independientes aleatorios, normales con media cero y matriz de dispersión constante.

A continuación se presentan las ideas centrales que sobre el análisis factorial ha desarrollado R.B. Cattell.

RAYMOND B. CATTELL

Cattell estudió con Spearman. En sus investigaciones utilizó el método de diseño multivariado, señalando que el análisis factorial produce un conjunto de factores llamados *rasgos* y conociendo su estructura en una persona, se pueden hacer predicciones acerca de cuál será su comportamiento en cualquier situación específica.

Para Cattell el objetivo principal del análisis factorial es representar o explicar relaciones observadas entre muchas variables experimentales en términos de dependencias lineales dentro y entre, reduciendo el número de variables conceptuales llamadas: "ideales", "abstractas" o "latentes". Ya que la substitución a un número reducido de conceptos y relaciones para un gran conjunto de particularidades es el objetivo de toda ciencia, el análisis factorial es uno de los métodos más directos, universalmente aplicable y representativo de los métodos científicos. El análisis factorial tiene las siguientes aplicaciones: 1) donde el número de variables a ser observado sobre un fenómeno en particular es muy grande, 2) donde se han tenido pequeños logros o éxitos, después de algunos años de alcanzar un acuerdo en los conceptos principales, y 3) donde existen buenas razones para esperar interacciones complejas, que no son fácilmente separables, experimentalmente hablando vía manipulación o control, de manera más notable cuando un número mayor de fenómenos está determinado de manera natural y múltiple.

Cattell reconoce que el Análisis Factorial tiene aplicaciones exploratorias y confirmatorias (prueba de hipótesis de estructuras factoriales dadas). Señala igualmente las proposiciones matemáticas básicas así como sus respectivas formulaciones, diciendo que típicamente el experimentador medirá n variables manifiestas u observadas a través de las respuestas a ciertos atributos en N sujetos u organismos (N > n).

En busca de posibles constructos de las correlaciones observadas, los conceptos psicológicos se relacionan entre n variables manifiestas que pueden ser representadas alternativamente como generadas por las relaciones de n variables a k factores.

DAVID J. BARTHOLOMEW

Para Bartholomew, el análisis factorial es un miembro de la familia de los métodos multivariados que también incluyen análisis de estructuras y de rasgos latentes. El común denominador de los modelos subyacentes a estos métodos es que las variables aleatorias observadas se asumen dependientes de variables aleatorias que son no-observables, es decir, variables latentes. Hay algunas discuciones acerca de si estas variables latentes son "*reales*" en cualquier sentido, pero ellas pueden considerarse simplemente como constructos diseñados para simplificar y resumir la compleja "*red*" de variables relacionadas con que se confrontan de manera natural. Para expresar estas relaciones en términos de un número pequeño de variables latentes, o factores, los modelos efectúan una reducción en dimensiones que ayuden a su comprensión.

Bartholomew intenta proporcionar un trabajo coherente dentro de la metodología existente el cual puede ser evaluado, así como una base para el desarrollo de nuevos métodos. Básicamente se destaca la discusión en términos de las distribuciones de las variables aleatorias involucradas y entonces hacer uso de las ideas de simetría, invarianza, independecia condicional, entre otras, con el objeto de determinar la clase de modelos razonables que serán considerados. Esto no solamente exhibe la unidad de varios modelos que ya existen, sino que resuelve muchas de las ambiguedades e incomprensibilidades con que el análisis factorial ha sido satanizado. Bartholomew (1980) cita a Lord & Novick (1968) quienes han trabajado en esta línea al igual que él también lo ha hecho.

La principal conclusión es que los métodos de variables latentes pueden proporcionar bases teóricas suficientes con una alta confiabilidad para justificar su uso como técnicas exploratorias. Al mismo tiempo algunas prácticas comunes, especialmente las relacionadas con los puntajes factoriales (*factor scores*) muestran la dependencia en supuestos de distribución no establecidos.

Teoría Básica

La clase de modelos que considera involucran 2 conjuntos de variables aleatorias. Las variables observables o manifiestas de dimensión p será representada por el vector x, las variables no observables o latentes de dimensión q por y o por z. No hay restricciones en cuanto a que las x's pueden ser discretas, contínuas, categóricas o mixtas. Su distribución conjunta será denotada por $f(x)$, que será interpretada como una densidad o función de probabilidad como la naturaleza de las variables. Las y's se supondrán como variables contínuas con densidad previa $h(y)$. Las cuales son consideradas como variables aleatorias y no como parámetros como en algunos tratamientos, porque la única característica que distingue a ellos de las x's es su falta de observabilidad, esto no se considera como una razón suficiente para darle un *status* o lugar diferente en la teoría. La razón de la restricción de que las y's deben ser contínuas es que un mayor número de variables latentes que ocurren en el trabajo práctico tienen este carácter.

Un modelo de variables latentes relaciona a las y's con las x's y de este modo nos permite hacer inferencias acerca de su distribución conjunta después de que las x's han sido observadas. Aunque q, la dimensión del espacio y, puede tomar cualquier valor, un modelo sólo será útil en la práctica si q es mucho menor que p. El objetivo principal del análisis basado en este modelo es doble:

a) efectúa una reducción en la dimensionalidad del problema expresando la estructura escencial de los datos en términos de las y's en vez de las x's.

b) localiza miembros muestreados en el espacio y en las bases de sus valores x observados. Esto es conocido como un problema de conglomeración. Un objetivo más que figura de manera prominente en la práctica es la interpretación de los "*Factores*", esto es, con identificación de las y's individuales con variables sustantivas latentes que son prácticamente significativas.

Bhartholomew dá un paso más al proponer el uso del término *"componente"* desde otro ángulo diferente al tradicionalmente empleado en componentes principales, definiéndolo como: *componente = factor + error*; de tal suerte que estas *componentes* pueden emplearse para examinar el signo y la magnitud de los coeficientes de las variables manifiestas, exactamente como en componentes principales a partir de la siguiente expresión:

λ_{ij}/ψ_i donde λ_{ij} son las cargas factoriales , i=1, 2,...,p j=1,2,...k y ψ_i son las varianzas únicas (*uniqueness*) de los factores únicos Bartholomew D. (1985, pp. 3,5).

KARL G. JÖRESKOG

En su libro "*Statistical Estimation in Factor Analysis*" (1963) señala que el Análisis Factorial es el término común dado a un número de técnicas estadísticas para la resolución de un conjunto de variables en términos de un número pequeño de variables *hipotéticas*, llamados *factores*. Esta resolución se lleva a cabo por medio de un análisis de las correlaciones de las variables. Si el análisis se realiza correctamente, la solución factorial puede producir factores que contengan toda la información esencial de las variables originales. Así, el principal objetivo es alcanzar una descripción llamada por Jöreskog *parsimoniosamente científica.*

Jöreskog distingue el modelo de componentes principales del modelo de análisis factorial a partir de las siguientes características distintivas: el modelo de componentes principales fue diseñado para determinar las componentes que presentan una mayor variabilidad, mientras que los modelos de análisis factorial estan destinados fundamentalmente a buscar los factores que contribuyen a explicar fuertemente a las correlaciones. Otra diferencia es el factor específico que se incluye en los modelos de analísis factorial pero no en el modelo de componentes principales. Los diversos modelos de análisis factorial están diferenciados sobre la base de proponer un factor general, la aprobación o aceptación de la necesidad de efectur la rotación y el criterio para la solución final (factores ortogonales vs. factores oblicuos, estructura simple, etc.). Afirmaciones apropiadas para un modelo a menudo no se aplican para otro modelo, aunque ciertas transformaciones que se le hacen a un modelo pueden hacersele a otro.

Particularmente lo que presenta a continuación Jöreskog trata con el modelo formulado por Thurstone (1947). La ecuación fundamental de este modelo expresa las variables observadas (manifiestas) como funciones lineales de factores hipotéticos (latentes) más un residuo que es un error de medición, tal y como se presenta en la ecuación (1).

Suponemos que este modelo sostiene que para una población específica de individuos, los datos con que se trabaja son representativos de una muestra de individuos extraída de esta población. Las cargas factoriales (que son los pesos en las combinaciones lineales) y las varianzas residuales son los parámetros del modelo. La especificación para este modelo incluye los problemas de: i) la existencia del modelo; ii) identificación de los parámetros y iii) determinación de los parámetros.

Estas consideraciones enfatizan la necesidad de imponer restricciones adicionales en los parámetros para hacer el modelo más determinado. La restricción es que las varianzas residuales sean proporcionales a los valores recíprocos de los elementos de la diagonal de la matriz de dispersión de la población. Este supuesto tiene ciertas relaciones con la teoría de la imagen de Guttman (1953). Otra ventaja es que el supuesto acerca de las varianzas residuales elimina el problema de la comunalidad. Si son necesarios pueden obtenerse como un resultado del procedimiento.

El uso más familiar de este modelo se encuentra en el análisis de pruebas mentales. Las variables latentes φ representan rasgos de habilidades humanas mentales. Los valores de φ en una población de individuos están distribuidos en k dimensiones. Cada componente de φ tiene una media igual a cero en esta población. Este supuesto sobre la media no constituye un camino restringido a la generalidad. Ya que φ no puede ser medido directamente, recurriremos a un conjunto de variables contínuas x cada una de las cuales está relacionada a φ , variando en diferentes grados y cada uno de los cuales es observable. Las combinaciones lineales de factores mentales forman parte de los puntajes de la prueba. Los renglones de Λ representan estas combinaciones lineales. Un componente ε constituye la parte del puntaje no *"explicado"* por los factores comunes φ. Lo cual se considerado integrado por un error de medición en el puntaje y un *factor específico.* Este factor específico actúa parcialmente y solamente con la prueba en particular. Si sólo un conjunto de observaciones en cada individuo está disponible, no podemos distinguir entre estas dos partes de ε. Podemos emplear, en este contexto, el término *residuo* para un componente de ε, lo cual implica que los residuos no están correlacionados. Esto es así por definición, es decir, las relaciones de los puntajes se consideran solamente producidos por factores comunes, y justamente esto es lo que se quiere investigar.

La expresión (1) es la ecuación fundamental en el modelo del análisis factorial múltiple de Thurstone [Thurstone (1947) pp. 63]. Estandariza las variables manifiestas x a varianzas unitarias, Σ es entonces una matriz de correlación. En su mayor parte los psicólogos formulan el modelo de análisis factorial en términos de correlaciones, debido a que la estandarización de los puntajes es un punto cuestionable por algunos investigadores, [Jöreskog (1963) sugiere revisar a Rasch (1953), Bartlett (1951) y Lawley (1956)] pudiéndose emplear la covarianza en lugar de las correlaciones.

La especificación del modelo de análisis factorial múltiple (MFA) como le llama Jöreskog es similar al análisis de regresión en donde cada componente de x es una combinación lineal de los componentes de φ más un residuo. La diferencia es que φ, que juega el papel de la variable explicativa (causal o independiente), es observable en el análisis de regresión pero no en el análisis factorial.

Los elementos de las matrices Λ, Φ y Ψ son los parámetros en el modelo. Estos son *parámetros estructurales* en el sentido que ellos afectan las variables manifiestas para todos los individuos. En el caso de que los elementos de Φ no sean aleatorios, decimos que $\varphi=\varphi_\alpha$ para el α-iésimo individuo, φ_α es considerado como un vector-parámetro caracterizando al α-ésimo individuo. Los φ_α son los *parámetros incidentales* en el sentido que sólo afectan a las variables manifiestas para el α ésimo individuo. Por lo tanto la ecuación (1) puede reescribirse como:

$$x_\alpha = \mu + \Lambda\varphi_\alpha + \varepsilon \tag{2}$$

donde

$$E(x_\alpha) = \mu + \Lambda\varphi_\alpha \tag{3}$$

y φ_α satisface (1). En este caso (3) es la media condicional de x para una φ dada. La distinción entre φ aleatorio y no aleatorio es la misma que se hace en el análisis de regresión y el análisis de varianza [(para mayor detalle Jöreskog (1963) sugiere ver a Graybill (1961), Cap. 5].

Para el modelo MFA distinguiremos los siguientes problemas estadísticos:

(i) Existencia e identificación de la estructura.

(ii) Estimación de los parámetros.

(iii) Estimación de los puntajes factoriales de los individuos.

(iv) Prueba de la hipótesis de que el modelo ajusta a los datos observados.

(v) Determinación del número de factores comunes.

El problema (i) es escencialmente un problema referente a la población. Los otros son problemas de inferencia estadística que se consideran una vez que se han hecho muestras observacionales. Pero como podemos ver, no es posible esperar que se haga una buena inferencia estadística de los datos a menos que tengamos una respuesta satisfactoria para el problema de identificación.

Identificación y existencia de la estructura.

A continuación se presentan algunos resultados del análisis factorial muy conocidos. Si se quiere un mayor detalle respecto de los mismos Jöreskog (1963) sugiere revisar a Reiersφl (1950 y Anderson & Rubin (1956).

Tomando las *esperanzas* de ambos lados de la ecuación (1) y empleando la ecuación (2) tenemos:

$$E(x) = \mu \qquad (4)$$

Empleando la ecuación (1) y el supuesto de que φ y e son independientes tenemos :

$$\begin{aligned}\Sigma = E[(x-\mu)(x-\mu)'] &= E[(\Lambda\varphi + e)(\Lambda\varphi + e)'] \\ &= E(\Lambda\varphi\varphi'\Lambda' + \Lambda\varphi e' + e\varphi'\Lambda' + ee') \\ &= \Lambda E(\varphi\varphi')\Lambda' + E(ee') \\ &= \Lambda\Phi\Lambda' + \psi \qquad (5)\end{aligned}$$

Este resultado final se debe a que los términos 2 y 3 de esta expresión son iguales a cero ya que anteriormente se señalo y y e no están correlacionados.

Esta ecuación dá una descomposición de Σ en 2 partes. La primera parte $\Lambda\Phi\Lambda'$ es positiva semidefinida y de rango k. Es la matriz de dispersión de las partes "comunes" de los puntajes y tiene exactamente los mismos elementos fuera de la diagonal como Σ. Los elementos de la diagonal de $\Lambda\Phi\Lambda'$ son llamados *comunalidades*. Thurstone (1947) emplea este término después de estandarizar los puntajes a varianza unitaria. Emplearemos el mismo término también para el caso no estandarizado. Frecuentemente, se hace el supuesto de que el vector x tendrá una distribución normal multivariada con vector media μ y la matriz de dispersión Σ que es de la forma presentada en la ecuación (2).

Jöreskog (1963) señala que siguiendo la terminología sugerida por Hurwics (1950) y también por Koopmans y Reiersφl (1950) podemos definir una *estructura* para ser considerada un modelo en el que todos los parámetros dados son numéricos.

Así, una estructura es una realización (representación) particular del modelo, y él mismo es el conjunto de todas las estructuras que satisfacen los postulados dados. Una estructura dada S= {Λ, Φ, Ψ} genera una y sólo una distribución de probabilidad conjunta P(x) de las variables manifiestas. Por otro lado, puede posiblemente haber algunas estructuras generadas por la misma distribución P(x). Si dos o más estructuras generan la misma distribución de probabilidad conjunta de las variables manifiestas, las estructuras se dice que son *equivalentes*. Si un parámetro tiene el mismo valor en todas las estructuras equivalentes, se dice que es *identificable*. Si un parámetro no es identificable existirá la estimación no consistente. El problema de existencia de una estructura radica ahora escencialmente en la pregunta siguiente: ¿ para una Σ dada, los supuestos del MFA y la ecuación (2) pueden encontarse *satisfactoriamente* las matrices Λ, Φ y Ψ ?. La respuesta a esta pregunta depende en cierto grado del valor del número k. Para algunos valores de k (por ejemplo $k = p-1$) la solución es siempre posible, mientras que para otros valores puede no existir una solución.

Si para algunos valores de k existe una solución habrá también infinitamente muchas otras soluciones con el mismo valor k y el mismo Ψ. Si $\{\Lambda,\Phi,\Psi\}$ es un conjunto de soluciones para un cierto valor de k y T es cualquier matriz kxk no singular, entonces $\{\Lambda^*,\Phi^*,\Psi\}$ es también un conjunto de soluciones con el mismo k, donde:

$$\Lambda^* = \Lambda T^{-1} \quad (6)$$

y $$\Phi^* = T\varphi T' \quad (7)$$

también notamos que si $\varphi^* = T\varphi$ (8)

entonces $\Lambda^*\varphi^* = \Lambda\varphi$

y (I) es $$x = \mu + \Lambda^*\varphi^* + e \quad (9)$$

la expresión (8) representa una transformación lineal no singular de los puntajes factoriales y (6) es la transformación inversa aplicada a los renglones de Λ. Todas estas transformaciones son equivalentes. Observando a x no podemos distinguir entre los modelos (1) y (9). Esto es una indeterminación fundamental del modelo. Algunas de estas no determinaciones pueden ser eliminadas si $\Phi = I$. En este caso se dice que los factores son *ortogonales*; si Φ no es diagonal los factores son *oblicuos*. Cuando asumimos que Φ=Y, T puede ser una matriz ortogonal si $\{\Lambda,Y,\Psi\}$ y $\{\Lambda^*,Y,\Psi\}$ serán dos estructuras equivalentes con el mismo valor de k. En este caso Λ es solamente determinado haciendo la post-multiplicación por una matriz ortogonal.

Aunque Λ^* y Λ son equivalentes desde el punto de vista matemático, desde el punto de vista psicológico no son equivalentes. El problema de escoger una Λ particularmente *significativa* desde el punto de vista psicológico fuera del grupo infinito $\{\Lambda T^{-1}\}$ se ha llamado el *problema de rotación*[1]. El que la rotación cuente con un fuerte apoyo entre los psicólogos quizás sea correcto. Pero otras veces el significado psicológico se traslada a destinar algunas restricciones establecidas precisamente en Λ, entonces es absolutamente concebible que la elección se convierte en un problema de estimación estadística.

J. P. BENZÉCRI

Benzécri y sus colaboradores han desarrollado una nueva *filosofía* para el análisis de datos, situándolo en el contexto general de la estadística, particularmente en el campo de las ciencias humanas, formulando y desarrollando el *Análisis de Correspondencias* como técnica central de dichos análisis. Este nuevo enfoque rompe con el paradigma clásico de la metodología estadística, proponiendo una serie de cambios. Los principios fundamentales de esta nueva disciplina se exponen muy brevemente a continuación (Hérnandez C. R. 1987 pp.445-446):

[1] El problema de transformación puede ser un término mejor, ya que incluye también la transformación de factores oblicuos, en cuyo caso la matriz de transformación T es no ortogonal y por lo tanto no representa sólo una rotación.

1).-*La estadística no es probabilidad.* Se propone el uso de un análisis estadístico independientemente de interpretaciones particulares de la probabilidad pero que está basado en el álgebra y en la geometría.

2).-*El modelo debe seguir a los datos y no al revés.* Este principio se refiere al hecho de que existe una tendencia a querer sobre todo *ajustar* modelos a los datos más que tratar de *extraer* estructuras de ellos.

3).-*Es conveniente tratar simultáneamente informaciones que conciernen al mayor número posible de dimensiones.* Aquí se comenta la diferencia entre *experimentación y observación* en Estadística. Sólo por medio del estudio sistemático de las relaciones entre variables observadas es posible proponer hipótesis. Este estudio implica el considerar el mayor número de variables posibles.

4).-*Para el análisis de hechos complejos y principalmente de hechos sociales, la computadora es indispensable.* Con un mayor número de información es necesario instrumentar nuevos métodos de análisis y síntesis.

5).-*Utilizar una computadora implica abandonar todas las técnicas concebidas antes de la llegada de la computadora. Técnicas, no ciencias.* Con el uso de dispositivos electrónicos (computadoras) se impone la revisión de todos los métodos estadísticos, lo que implica el fin de no pocas técnicas.

El *análisis de datos* puede ser entendido como un conjunto de técnicas fundamentalmente descriptivas basadas en el álgebra y en la geometría que permiten el estudio de grandes volúmenes de datos multidimensionales de todo tipo (numéricos o no) por medio del uso intensivo de programas de cálculo electrónico. Cabe subrayar que en este conjunto de técnicas, al menos en su origen, no tiene carácter inferencial en el sentido clásico de la estadística por lo que debe tenerse presente esta diferencia al tratar de comparar el análisis de datos con otras escuelas estadísticas, particularmente refiriéndose sólo al análisis de datos de Benzécri.

Otro aspecto que resulta interesante destacar en esta nueva disciplina del análisis de datos es que el énfasis está puesto no en el aspecto del modelo probabilístico que subyace a las técnicas multivariadas, como la densidad normal multivariada que debe tener la población de donde se extrajeron los datos, sino en la evolución del concepto multivariado.

La generalización del análisis de correspondencias para cualquier número de criterios nominales y con la existencia de un buen acervo de programas de cómputo, realiza todos los análisis de datos, de manera eficiente. Observándose tres características importantes que se desarrollaron en la evolución del análisis de datos:

1) El fin del aislamiento que había mantenido al análisis de datos como una "escuela francesa de estadística".

2) La aparición de las computadoras personales vino a dar un nuevo impulso al análisis de datos. Los paquetes que se han desarrollado para tal fin han permitido que el análisis de datos se utilice en nuevas áreas.

3) El análisis de correspondencias, en su práctica actual, ha tenido una preocupación acerca de la estabilidad en sus resultados. Esto ha traído como consecuencia el uso de técnicas de remuestreo, el empezar a combinar de manera explícita técnicas del análisis de datos de la estadística clásica e inferencial. Para mayor detalle de los trabajos realizados se sugiere revisar las obras citadas por Hérnández C. R. (1987).

En conclusión Hérnández *et al* concluye su trabajo diciendo que el análisis de datos aparece cada vez más como un grupo de técnicas intermedias entre lo exclusivamente descriptivo y lo exclusivamente inferencial.

Finalmente estas técnicas particularmente en el campo de las ciencias sociales, permiten una mayor flexibilidad en su uso y en consecuencia un avance en el conocimiento de éstas áreas, ante la dificultad que existe no sólo para medir ciertos aspectos de la propia conducta sino de *ajustarla* a los supuestos que subyacen al uso de las técnicas estadísticas *clásicas* o *tradicionales*.

CAPITULO 2

ANALISIS FACTORIAL

El término Análisis Factorial es el término que se emplea para designar al conjunto de técnicas estadísticas de análisis multivariado, que tienen como principal propósito explicar numerosas variables llamadas manifiestas u observables por medio de un grupo menor de variables llamadas latentes o simplemente factores.

Aunque el análisis factorial ha tenido su principal desarrollo y aplicación en el campo de la Psicología, particularmente en la construcción de instrumentos (valídez y confiabilidad) también se aplica en otras áreas del conocimiento, tales como, Meteorología, Ciencia Política, Medicina, Geografía, Administración y Economía, entre otras (ver e.g Reyment Richard y K.G. Jöreskog, 1993).

El análisis factorial puede formularse en términos de correlaciones parciales, entre las variables manifiestas $\boldsymbol{x}' = (x_1, \ldots, x_p)$ de p variables y las variables latentes $\boldsymbol{y}' = (y_1, \ldots, y_k)$ de k variables contínuas. Justamente intenta "explicar" estas correlaciones mediante un análisis que produce una estructura subyacente integrada por un número menor de factores que contienen toda la información escencial acerca de las relaciones lineales entre los valores de las variables manifiestas u observables $\boldsymbol{x}'$.

Especialmente en Psicología los puntajes con los cuales se trabaja son estandarizados con media igual a cero y desviación estándar igual a uno y las correlaciones a las que anteriormente se ha hecho referencia para este tipo de variables contínuas son las correlaciones poblacionales producto-momento de Pearson denotadas por ρ_{xixj} y calculada a partir de la siguiente expresión matemática:

$$\rho_{xixj} = \sum x_i x_j / N\sigma_{xi}\sigma_{xj}$$

donde:

$$\sigma_{xi} = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \bar{x}_i)^2}{N}} \quad \sigma_{xj} = \sqrt{\frac{\sum_{i=1}^{N}(x_j - \bar{x}_j)^2}{N}}$$

$$\bar{x}_i = \sum_{i=1}^{N} x_i / N \quad \bar{x}_j = \sum_{j=1}^{N} x_j / N$$

N=número de elementos muestrales.

Es conveniente hacer la aclaración que también puede emplearse otro tipo de correlaciones como la correlación tetracórica inicialmente empleada por Thurstone, simbolizada por: r_t (frecuentemente usada en Psicología) para el caso en el cual las variables manifiestas han sido reducidas artificialmente o toman sólo, 2 valores. La correlación tetracórica es un caso general de la correlación policórica cuando se tienen más de dos valores. Suponiendo que z_1 y z_2 son dos variables a nivel ordinal con variables subyacentes z^*_1 y z^*_2, respectivamente y asumiendo que estas tienen una distribución normal bivariada, su correlación se conoce como *policórica.*

Cuando hay que calcular un gran número de correlaciones tetracóricas r_t, los diagramas de cálculo de Thurstone pueden aportar una considerable economía de trabajo cuando las operaciones se hacen manualmente.

Vale la pena mencionar que en el análisis factorial las variables $x' = (x_1, \ldots, x_p)$ son variables contínuas observables o manifiestas y existen modelos en donde puede haber una gran mayoría de ellas y algunas dicotómicas, donde justamente puede emplearse la correlación producto momento de Pearson, pero cuando hay una mezcla de variables a nivel ordinal e intervalar en proporciones similares lo adecuado es emplear la correlación policórica.

Como podemos observar de lo anteriormente expuesto, el análisis factorial parte de una técnica estadística: las correlaciones lineales de las variables manifiestas u observables x'. El análisis factorial se propone explicar estas correlaciones introduciendo factores subyacentes denotados por $y_1, y_2, \ldots y_k$ los cuales son explicados por dichas correlaciones.

Consideramos que el aspecto más valioso del análisis factorial estriba en el hecho de "descubrir" estructuras subyacentes de fenómenos complejos tal y como son los que tienen que ver con el comportamiento humano. Ahora bien la pregunta que surge de manera natural es la siguiente: ¿cómo y qué hace el análisis factorial para conseguirlo? . Para responder esta pregunta es necesario comprender algunos conceptos estadísticos básicos que serán expuestos a continuación.

En primer término suponemos que las variables x y y tienen una distribución normal multivariada (no necesariamente). Entonces la distribución conjunta $f(x, y)$ y la distribución condicional $g(x|y)$ son ambas también normales. Además la media de x dado y es lineal en y y la matríz de dispersión condicional no depende de y. Teniendo entonces el análisis factorial el modelo siguiente (que es el mismo propuesto por Thurstone):

$$x = \mu + \Lambda y + e \qquad (1)$$

donde:

x = vector de variables manifiesta u observadas.

Λ= matriz de cargas factoriales

y = vector de variables latentes, también llamadas factores.

e =vector de variables específicas (residuos), de elementos independientes aleatorios, normales con media cero y matriz de dispersión constante.

Note que el no incluir a μ es irrelevante al modelo, ya que μ es una constante que representa a la media.

En forma de ecuaciones lineales sería:

$$
\begin{aligned}
x_1 &= \lambda_{11}y_1 + \lambda_{12}y_2 + \ldots \lambda_{ik}y_k + e_1 \\
x_2 &= \lambda_{21}y_1 + \lambda_{22}y_2 + \ldots \lambda_{2k}y_k + e_2 \\
&\vdots \qquad\qquad\qquad\qquad\qquad\qquad (2) \\
x_p &= \lambda_{p1}y_1 + \lambda_{p2}y_2 + \ldots \lambda_{pk}y_k + e_p
\end{aligned}
$$

Podemos observar claramente estas ecuaciones (2) representan lo que Jöreskog (1979, p. 7) llama *ecuaciones de regresión* lineales por mínimos cuadrados de una variable en los factores y_1, $y_2, \ldots y_k$. Una propiedad bien conocida de la regresión por mínimos cuadrados es que los residuos e_i en la i-ésima ecuación no están correlacionados con los e_j's, esto es, la corr(e_i, e_j)=0 , $i \neq j$

El objetivo principal de este modelo de análisis factorial es determinar el valor de k, es decir, el número de variables latentes y los elementos de las matrices Λ y Ψ, dada $S = (s_{ij})$ la matriz de covarianza muestral de Σ.

donde

$$\Sigma = \Lambda\Phi\Lambda' + \Psi \qquad (3)$$

Si Φ=I y suponiendo que los factores no están correlacionados entonces la ecuación (3) será igual a $\Sigma = \Lambda\Lambda' + \Psi$;

$$
E(ee') = \Psi = \begin{pmatrix} \psi_1^2 & 0 & \ldots & 0 \\ 0 & \psi_2^2 & \ldots & 0 \\ 0 & 0 & \ldots & \psi_p^2 \end{pmatrix}
$$

esta es la matriz de covarianza del vector aleatorio de residuos e y además cov(e,y) = 0

Antes de continuar esta exposición, es conveniente aclarar el término de *no determinación* (indeterminancia) mencionado anteriormente Un modelo factorial no puede identificarse hasta que el número de los factores se ha establecido. Si esto no se establece, existe indeterminancia entre la varianza de un factor y las cargas de la variables observadas o manifiestas en el factor. Esto hace imposible distinguir entre el caso en que un factor tiene una varianza grande y sus cargas factoriales son pequeñas, y el caso en que la varianza es pequeña y las cargas en el factor son grandes.

En general si las varianzas de los factores se fijan (o bien algunas cargas factoriales) en cada factor, la no determinación de k se elimina del modelo factorial. Esto no necesariamente significa que el modelo esté identificado, pueden aún existir otras fuentes de no determinación de k. Este problema se resolvió con los trabajos realizados por Jöreskog (1973, 1979) como fue señalado en el capítulo referente a los antecedentes históricos del análisis factorial. Conjuntando el análisis de la estructura de la covarianza y el modelaje de ecuaciones lineales estructurales en un sólo modelo general llamado LISREL (*Linear Structural Equation Model for Latent Variables*) .Estos modelos son mucho más generales y no se presentarán en este trabajo (ver, Jöreskog y Sörbom 1979 además de Loehlim, John C., 1992)

Suponiendo que k se conoce y por ello se fija, entonces con $\frac{1}{2}p(p+1)$ elementos de S, es necesario estimar $p(k+1)$ parámetros desconocidos. Haciendo una transformación de un conjunto de factores y a otro lo que corresponde a una transformación o rotación de ejes coordenados en el espacio euclidiano, las y's dejarán de ser no determinadas, estimando para ello el espacio factorial (Factor Space), ¿cómo se hace esto?, determinando un único conjunto de ejes en donde el espacio esté sujeto a la condición siguiente:

$$\Gamma\Gamma' = \Lambda \qquad (4)$$

donde Γ es una matriz ortogonal[1] de orden (kxk) y Λ es una matriz diagonal, cuyos elementos en la diagonal son λ_i. Esto nos permite tener la libertad de rotar los ejes en cualquier posición que nos proporcione una significativa o mejor interpretación de las cargas o ponderaciones factoriales. Una transformación ortogonal será aquella en que los ejes son ortogonales en el sentido de que los factores no estan correlacionados, dando el ángulo en los 2 ejes coordenados (factores) de 90° y por lo tanto su correlación es igual a cero. Por convención los ejes se rotan en la forma que lo hacen las manecillas del reloj, manteniendo la ortogonalidad entre ellos. Si los factores rotados están correlacionados la rotación se llama oblicua y hay diferentes aproximaciones a la misma, tales como : 1) el método de ejes de referencia y 2) el método del empleo de la estructura o matriz patrón, más adelante se verán con mayor detalle las rotaciones factoriales.

Dos principios básicos en el Análisis Factorial son:
1) Principio de Independencia Lineal Condicional
2) Principio de estructura simple (Parsimonia)

El primero de ellos se refiere al hecho de que los factores pueden explicar todas las correlaciones lineales entre las variables. Una vez que los factores han sido parcialmente eliminados, quedan variables no correlacionadas y en este sentido el análisis factorial es un método de *clasificación de dependencia lineal.* Habrá un número de factores de grado k en p cuando se ha encontrado algun $k<p$ factores $y_1, y_2, \ldots, y_k$ tal que:

$$\rho(x_i, x_j\,; y_1, y_2, \ldots, y_k) = 0, \qquad i \neq j \qquad (5)$$

en cuyo caso tenemos una dependencia lineal de grado k entre las variables manifiestas donde:

$$\rho(e_i, e_j) = 0 \quad , i \neq j$$

Cada ecuación de (3) representa la ecuación lineal por mínimos cuadrados de una variable en los factores $y_1, y_2, \ldots, y_k$. Los coeficientes λ_{ik}, son coeficientes y llamados en Psicología *cargas factoriales.*

[1] Una matriz cuadrada Γ es ortogonal si $\Gamma\Gamma' = I$, donde I, es la matriz identidad definida $I=(c_{ij})$ donde $c_{ij} = \{1$ si $i=j$ y 0 si $i \neq j\}$.

Los factores $y_1, y_2, \ldots y_k$ son llamados *factores comunes* o simplemente *factores* porque miden atributos que son comunes a 2 ó más variables. Los residuos $e_1, e_2, \ldots, e_p$ se llaman residuos o *factores únicos*. Un factor único es usualmente considerado como una suma de *errores de medición* y un *factor específico*. Ni los errores de medición ni el factor específico contribuyen a cualquier correlación entre variables. Cada ecuación en (3) descompone a la variable x_i en 2 partes no correlacionadas.

$$x_i = c_i + e_i \qquad (6)$$

donde

$$c_i = \lambda_{i1}y_2 + \lambda_{i2}y_2 + \ldots + \lambda_{ik}y_k \qquad (7)$$

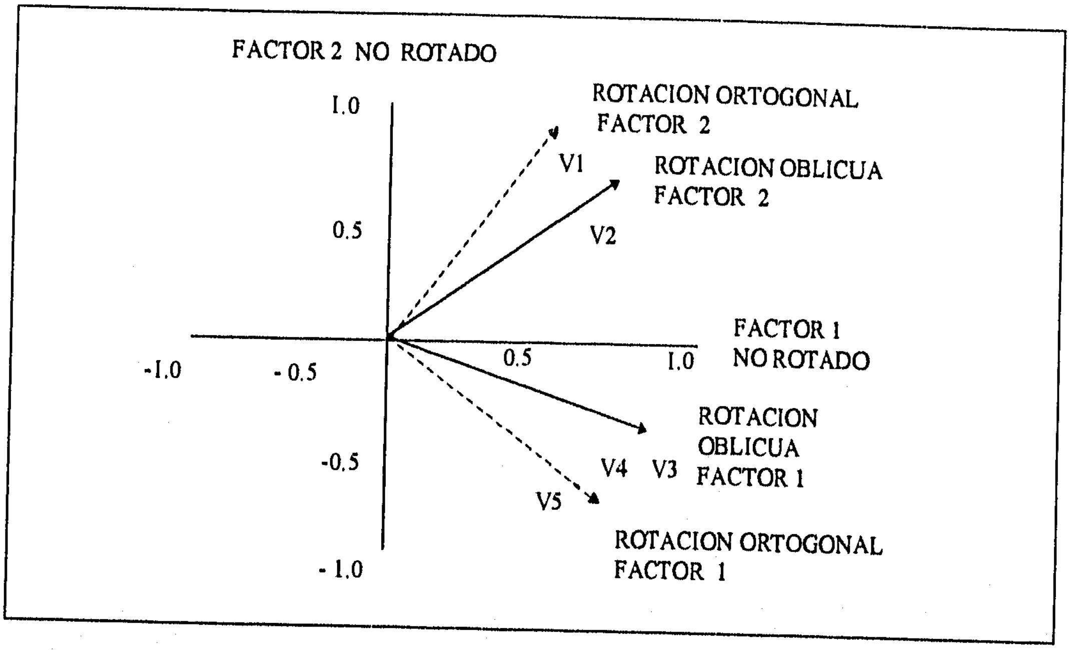
y se llama *parte común*. En otras palabras es la parte de cada variable que es común a algunas de las otras p-1 variables manifiestas y e_i es el residuo de cada variable que es único.

En cuanto al segundo principio del Análisis Factorial llamado de la estructura simple (Parsimonia) diremos que una vez que, un conjunto k factores estimados se ha encontrado para las correlaciones de las pruebas, las ecuaciones lineales (2) pueden transformarse o rotarse por otro conjunto de k factores que estimen igualmente bien a las correlaciones observadas. En efecto, cualquier transformación lineal no singular del primer conjunto de factores produce un nuevo conjunto de factores con esa propiedad. Sin embargo hay infinitos conjuntos de factores, esto es, una gran indeterminancia en el modelo.

Thurstone (1947, citado en Jöreskog 1979 p. 12) propuso como una solución a este problema, tomar en cuenta sólo a los factores para los cuales las variables tengan una simple representación. La matriz de cargas factoriales tendrán tantos elementos como sea posible con cargas igual a cero. Si $\lambda_{im} = 0$, el m-ésimo factor no se considera dentro de la i-ésima prueba. Una variable puede no depender en todos los factores comunes y sólo en una pequeña parte de ellos. También el mismo factor puede estar en sólo una pequeña parte de las variables. Tal matriz es considerada como la que proporciona la estructura simple y *presumiblemente* con la interpretación psicológica más útil, siendo este el concepto de *estructura simple*.

Desde el punto de vista de las Ciencias Sociales aplicadas tales como la Psicología, la mejor estructura simple generalmente se consigue con los factores oblicuos, ¿por qué?. En primer lugar porque los factores que así se generan son flexibles, ya que no tienen el supuesto de tener que ser ortogonales, esto es más realista, porque las dimensiones subyacentes teóricamente importantes en la realidad generalmente se dan correlacionados unos a otras. Inclusive también frecuentemente puede observarse una estructura más definida que la obtenida con la rotación ortogonal, en términos de que las variables con poca carga factorial en la rotación oblicua tienen valores aún mucho más pequeños, en muchas ocasiones con valores cercanos a cero como podremos apreciar en los ejemplos que se plantean más adelante. Por ejemplo presentamos la siguiente gráfica.

ROTACION FACTORIAL



Gráfica No. 1

A continuación se presentan los datos hipotéticos de la gráfica anterior.

DATOS HIPOTETICOS

	CARGAS FACTORIALES NO-ROTADAS		CARGAS FACTORIALES ROTADAS	
VARIABLES	FACTOR 1	FACTOR 2	FACTOR 1	FACTOR 2
V1	.50	.80	.03	.94
V2	.60	.70	.16	.90
V3	.90	-.25	.95	.24
V4	.80	-.30	.84	.15
V5	.60	-.50	.76	-.13

Tabla No.1

Esta precisión es el resultado del hecho de que cada eje factorial rotado es ahora cercano al grupo respectivo de variables. Los factores oblicuos nos proporcionan información acerca del alcance de las correlaciones entre los factores. Uno de los problemas técnicos que presenta este procedimiento analítico es que aún no está bien desarrollado y por lo mismo es sujeto de considerables controversias. Cuando el análisis factorial va a ser empleado en análisis estadísticos subsecuentes, sí se elige un procedimiento de rotación ortogonal se elimina con ello la posible colinealidad entre los factores. Igualmente es deseable emplear factores ortogonales cuando elinterés se centra en obtener constructos o dimensiones teóricamente significativos. Por todo lo anterior los factores oblicuos son en la práctica más realistas.

La implicación práctica de esta discusión se hizo con el objeto de determinar el número k de variables latentes que pueden incluirse en el modelo.

Tenemos $\frac{1}{2}p(p+1)$ elementos de S para estimar los $p(k+1)$ parámetros (pk cargas factoriales y p varianzas específicas), pero la ecuación (5) contiene $\frac{1}{2}k(k-1)$ restricciones (los elementos fuera de diagonal de la matriz simétrica $\Lambda\Lambda'$ todos son cero). Para la estimación de los parámetros, necesitamos $\frac{1}{2}p(p+1) \geq p(k+1) - \frac{1}{2}k(k-1)$, de modo que k debe satisfacer

$(p-k)^2 \geq p + k$.

Por ejemplo, con 5 variables manifiestas no podemos tener más de 2 variables latentes (factores), con 8 variables manifiestas no podemos tener más de 4 variables latentes (factores) y con 15 variables manifiestas no podemos tener más de 10 variables latentes (factores). En la práctica, no tenemos que quedarnos con los factores máximos de esta relación, sino con aquellos que nos permitan explicar lo mejor posible las correlaciones entre las variables manifiestas.

Considerando la ecuación (6), dado que la parte común c_i y e_i no están correlacionados, esta ecuación también particiona la varianza de x_i como:

$$\mathrm{var}(x_i) = \mathrm{var}(h_i) + \mathrm{var}(e_i) \qquad (8)$$

a la var(h_i) se le llama *varianza común* o *comunalidad* de x_i y a la var(e_i) se le llama *varianza única* (*unique variance*) o la (*especificidad*) de x_i. La comunalidad de una variable es la proporción que una variable contribuye a la explicación de la varianza total de los factores comunes.

La varianza única es la proporción no explicada por los factores comunes. Para determinar la comunalidad de x_i, empleamos la ecuación (7) y tenemos:

$$\begin{aligned} E(h_i)^2 = {} & \lambda^2_{i1}E(y^2_1) + \lambda^2_{i2}E(y^2_i) + \ldots + \lambda^2_{ik}E(y^2_k) \qquad (9) \\ & + 2\lambda_{i1}\lambda_{i2}E(y_1 y_2) + 2\lambda_{i1}\lambda_{i3}E(y_1 y_3) + \ldots \\ & + 2\lambda_{i,k-1}\lambda_{ik}E(y_{k-1}y_k) \end{aligned}$$

Si $y_1, y_2, \ldots, y_k$ tienen todos varianza igual a uno, esta ecuación se convierte en:

$$\begin{aligned} \mathrm{var}(h_i) = {} & \lambda^2_{y1} + \lambda^2_{y2} + \ldots + \lambda^2_{ik} \\ & + 2\lambda_{i1}\lambda_{i2}\rho(y_1,y_2) + 2\lambda_{i1}\lambda_{i3}\rho(y_1,y_3) + \ldots \\ & + 2\lambda_{i,k-1}\lambda_{jk}\rho(y_{k-1},y_k) \end{aligned} \qquad (10)$$

Si los factores no están correlacionados, la expresión anterior se reduce aún más:

$$\mathrm{var}(h_i) = \lambda_{i1}^2 + \lambda_{i2}^2 + \ldots + \lambda_{ik}^2 \qquad (11)$$

que es simplemente la suma de cuadrados de los elementos en el i-ésimo renglón de Λ. Al hacer rotaciones ortogonales a los factores, a las comunalidades y a los residuos o factores únicos no se ven afectados por dichas transformaciones, ya que los factores únicos son ellos mismos. Por lo tanto sus varianzas, y sus comunalidades son invariantes ante tales transformaciones.

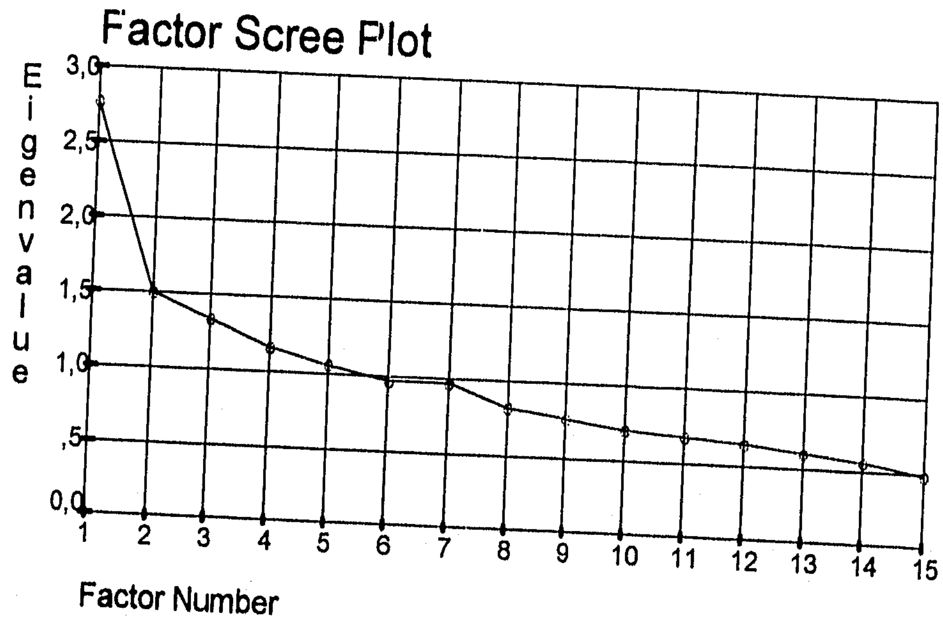
Además de esto existen algunos criterios adicionales con respecto al número de factores a considerarse, algunos de ellos son:

a) criterio de la raíz latente (eingenvalores). En análisis factorial el valor de los eigenvalores puede ser menor o aproximado a la estimación del promedio de las comunalidades estimadas para todas las variables.Este criterio de la raíz latente es confiable cuando el número de variables manifiestas se encuentra entre 20 y 50 variables, para variables < 20 el criterio es conservador y para >50 variables se pueden extraer muchos factores (Joseph F. Hair 1992, pp.239).

b) criterio a priori. Es útil si el investigador o analista está sometiendo a prueba una teoría o hipótesis acerca del número de factores considerados, o bien cuando se desea replicar trabajos.

c) criterio del porcentaje de varianza. Este criterio se basa en el porcentaje acumulado de la varianza extraída por factores sucesivos. No hay absolutamente ningun criterio que se haya adoptado, sin embargo algunos de ellos en la práctica se han dado. Para las ciencias sociales, donde a menudo la información que se obtiene es menos precisa, se considera apropiado un 60% del porcentaje de la varianza total.

d) criterio de la prueba de la gráfica de *"desmoronamiento"* (*Scree Test*). Este criterio fue estudiado y propuesto por Cattell. Como sabemos en el análisis factorial, todos los factores contienen al menos algo de varianza única, desde luego que la proporción de varianza única en estos últimos factores es substancialmente menor que en los primeros factores y justamente se retiene el valor de k a partir de que la gráfica no presenta cambios "sustanciales" en su comportamiento. Un ejemplo de este criterio se presenta en la gráfica siguiente.



Gráfica No. 2

El punto en que la curva tiende a *"enderesarse"* y se hace más estable, se considera como indicador del máximo número de factores que serán extraídos, que en esta gráfica hipotética serían 6 los factores.

Además de los criterios antes señalados, es importante tomar en cuenta la habilidad y conocimientos sobre la materia objeto de estudio para asignar algún significado útil a los factores interpretando la naturaleza de las variables, lo cual es muy importante en este proceso de extracción.

Algunos de estos criterios para la significación de las cargas factoriales son los siguientes (Joseph F. Hair op. cit., pp.239):

+ de .30 significativo
+ de .40 más significativo
> de .50 muy significativo

Esta clasificación es igual a las hechas por algunos investigadores en Psicología, tales como Kerlinger (1981, pp. 180) y Guilford (1984, pp. 398) entre otros.

ANALISIS FACTORIAL Vs. ANALISIS DE COMPONENTES PRINCIPALES.

Vale la pena dedicar algunas líneas de este trabajo a aclarar que el análisis factorial es diferente al análisis de componentes principales. Desafortunadamente el análisis factorial es frecuentemente confundido con el análisis de componentes principales, como si se tratara de lo mismo. Podríamos decir, que esta confusión no es gratuita pues los 2 métodos de análisis multivariados son similares en ciertos aspectos pero los resultados son totalmente diferentes. Particularmente Jöreskog (1979) cita a Kendall y Lawley (1956), Lawley y Maxwell (1971) quienes han enfatizado en tal distinción, así como otros.

Empecemos señalando las diferencias entre ambas técnicas estadísticas de análisis multivariado:

El análisis factorial se enfoca a explicar las *correlaciones* , en tanto que el análisis de componentes principales se dirige a explicar la *varianza* . Mientras que el análisis factorial se orienta a reproducir las correlaciones de las variables, el análisis de componentes principales se enfoca a reproducir la varianza total.

Aunque en el análisis de componentes principales, es posible que con unos cuantos componentes extraídos se explíque una proporción grande de la varianaza total, todos los componentes se requieren para reproducir exactamente las correlaciones originales. En tanto que en análisis factorial, hay por definición un cierto número de factores desde luego menores que el número de variables manifiestas originales, que reproducen las correlaciones exactamente. Sin embargo estos factores no explican una gran proporción de la varianza total como lo hace componentes principales con este mismo número de componentes.

Para ejemplificar lo anteriormente expuesto, vale la pena reproducir las tablas de un ejemplo tomado de Jöreskog (1979. pág. 13 y 16):

MATRIZ DE CORRELACION

	x_1	x_2	x_3	x_4	x_5	x_6
x_1	1.000	0.720	0.378	0.324	0.270	0.270
x_2		1.000	0.336	0.288	0.240	0.240
x_3			1.000	0.420	0.350	0.126
x_4				1.000	0.300	0.108
x_5					1.000	0.090
x_6						1.000

Tabla No. 2

Para estos datos, Jöreskog determinó dos factores que para el caso de análisis de componentes le llama componentes y para análisis factorial le llama factores con los resultados se presentan en la tabla siguiente.

COMPARACION ENTRE ANALISIS FACTORIAL Y COMPONENTES PRINCIPALES

Pruebas de habilidad	Análisis Factorial		Análisis de Componentes Principales *	
	Comunalidad	e_i	varianza explicada	varianza residual (v_i)
1	0.81	0.19	0.75	0.25
2	0.64	0.36	0.72	0.28
3	0.49	0.51	0.59	0.41
4	0.36	0.64	0.55	0.45
5	0.25	0.75	0.51	0.49
6	0.09	0.91	0.51	0.49
Total	2.64	3.36	**3.63**	**2.37**

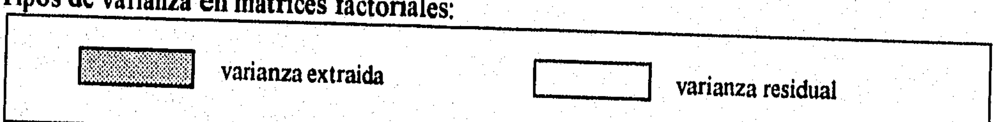
Tabla No. 3

* donde los componentes de varianza unitaria se estandarizaron con objeto de presentar una similaridad formal entre el anális factorial y el análisis de componentes.

Como podemos apreciar en esta tabla, el total de la varianza explicado por los 2 componentes es de 3.63 que en términos porcentuales es del 56% comparado con el 2.64 que equivale al 44% explicado por los dos factores, las varianzas residuales en componentes son menores que en análisis factorial, excepto para la primera prueba (x_1). La reflexión que estos valores nos merece es que, generalmente en el análisis factorial se incluye en cada prueba el factor específico e_i, no así en v_i , por ello en el análisis factorial, salvo para la primera prueba, todos los valores son mayores que v_i.

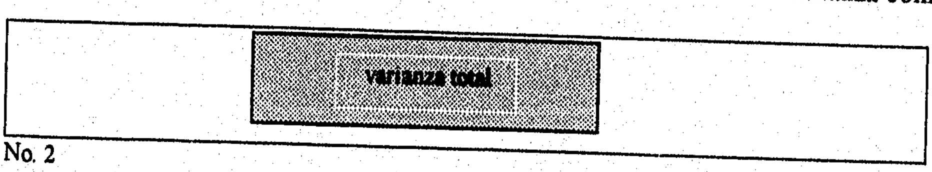
De manera esquemática podemos ilustrar lo anterior:

Tipos de varianza en matrices factoriales:



Cuadro No. 1

En componentes principales en la matriz de correlación (**R**) se tienen unos en la diagonal, por lo tanto la solución de los componentes obtenida se basa solamente en la varianza común.



Cuadro No. 2

En análisis factorial en la diagonal de la matriz de correlación se tienen los valores de las comunalidades y los factores que se obtienen se basan sólo en la varianza común.

varianza comun	varianza específica + error

Tabla No. 3

Regresando al ejemplo propuesto por Jöreskog (1979), de la misma forma como en análisis factorial, las dos componentes principales pueden transformarse linealmente en otro conjunto de componentes que, juntos expliquen tanto de la varianza como las componentes principales originales. En componentes principales se puede hacer para obtener alguna interpretación interesante, pero generalmente no la tiene, excepto en términos de las propiedades de sus varianzas y correlación (Jöreskog, 1979 p. 16).

Otra diferencia más entre estas dos técnicas estadísticas es que los componentes principales son por definición combinaciones lineales de variables $x_1, x_2, \ldots x_p$ mientras que los factores no lo son. Los factores son combinaciones lineales de las partes comunes $c_1, c_2, \ldots, c_p$ de las variables. Donde $c_i = \lambda_{i1}y + \lambda_{i2}y + \ldots + \lambda_{ik}y_k$ definidas en (7).

Jöreskog estableció para el ejemplo anterior la siguiente estructura:

$$\Lambda^{**} = \begin{pmatrix} 0.9 & 0 \\ 0.8 & 0 \\ 0 & 0.7 \\ 0 & 0.6 \\ 0 & 0.5 \\ 0.3 & 0 \end{pmatrix}$$

Por ejemplo, los factores y_1^{**} y , y_2^{**} tienen la siguiente representación:

$$y_1^{**} = (0.9^2 + 0.8^2 + 0.3^2)^{-1} (0.9c_1 + 0.8c_2 + 0.3c_6)$$
$$y_2^{**} = (0.7^2 + 0.6^2 + 0.5^2)^{-1} (0.7c_3 + 0.6c_4 + 0.5c_5)$$

Esto es fácilmente verificable substituyendo $c_1, c_2, \ldots, c_6$ por la expresión (7).

Componentes principales no es un modelo en el sentido usual de la palabra, es meramente un método descriptivo de análisis que justamente puede emplearse para analizar toda clase de variables cuantitativas. Análisis factorial sí es un modelo: $x = \Lambda y + e$ el cual está sujeto a prueba con datos empíricos. Las ecuaciones de este modelo no están sujetas a contrastación directa, debido a que p variables x_i son expresadas en términos de otras $p + k$ variables que no son observables. Pero las ecuaciones de este modelo si implican una hipótesis que puede ser probada, las correlaciones de las x's son de la forma siguiente:

$$\rho_{ij} = \lambda_{i1}\lambda_j + \lambda_{i2}\lambda_{j2} + \ldots + \lambda_{ik}\lambda_{jk} \qquad (12)$$

Si k es menor que p, entonces (12) impone restricciones en las ρ's, y este es el modelo más restrictivo que existe en el análisis factorial.

Hemos hecho referencia a lo que Jöreskog (1965) llamó ecuaciones lineales de regresión en esta primera parte, por lo cual es conveniente aclarar las diferencias entre el Análisis de Regresión y el Análisis Factorial como se hizo con el Análisis de Componentes Principales.

En el caso del análisis de regresión lineal, los λ's son los coeficientes de regresión los e's son los residuos que no están correlacionados con las y's, éstas sí se conocen y la única variable aleatoria es e y por lo tanto también la variable dependiente. En análisis factorial los λ's son los coeficientes de los factores los cuales se asume se determinan, los e son los residuos y no están correlacionados con los y's, esto es, con los factores, los y's y los e's son variables aleatorias y las y's no se conocen.

En resumen el análisis factorial es un modelo lineal basado en una técnica estadística que permite probar hipótesis sobre la estructura de *asociación* entre un conjunto de variables manifiestas. Si en el análisis factorial, las varianzas específicas son pequeñas, pueden obtenerse resultados similares con el análisis de componentes principales, pero si por el contrario las varianzas específicas son "grandes", el análisis de componentes principales las asimila independientemente de los factores y en consecuencia los dos análisis darán diferentes interpretaciones como ya se presentó anteriormente. En el análisis factorial, cuando se usa el método de estimación de máxima verosimilitud, hay una relación directa entre los resultados obtenidos de la matriz de correlación (esto es, datos estandarizados) y de la matriz de covarianza (esto es, datos crudos) no así en el análisis de componentes principales. Finalmente, la estimación de los puntajes factoriales requiere de operaciones iterativas, mientras que los puntajes en componentes principales se obtienen directamente. Por lo tanto, antes de elegir cuál de los dos análisis emplear, es necesario que el investigador tenga claro los objetivos que se consiguen con cada análisis y el uso que le dará a los resultados obtenidos

CAPITULO 3

METODOS DE ESTIMACION

Una vez que se ha identificado el modelo, se puede investigar, entre otras cosas, si los parámetros del modelo tienen alta probabilidad de ocurrencia, en este sentido, el objetivo es estimar los parámetros poblacionales subyacentes (baja la hipótesis dada) que tendrán la mayor verosímilitud de predecir la matriz de correlaciones observadas. Si un modelo no es identificado, es imposible determinar los parámetros, aunque los valores de cada variable observada sean conocidos para la población.

Para llevar a cabo la estimación además de que se SUPONE que el modelo se ha identificado ésta emplea los datos muestrales para estimar los parámetros poblacionales. El objetivo general de la estimación en el modelo factorial es encontrar estimadores de los parámetros que reproduzcan la matriz muestral de varianzas y covarianzas de las variables observadas tan cercanamente como sea posible en función de un criterio bien definido.

En el trabajo práctico con datos empíricos (muestrales) el modelo del análisis factorial no se ajusta "perfectamente" a los datos. Entoces es claro que el problema estadístico de estimación que tiene que resolverse es el siguiente: ¿Cómo ajustar convenientemente la matriz modelo Σ a partir de una matriz de varianza-covarianza muestral S?.

Para dar respuesta a esta pregunta es necesario definir algunos conceptos que veremos a continuación, los cuales nos permitirán integrarlos a los principales métodos de estimación desarrollados, para su mejor comprensión.

Recordemos que el modelo básico del análisis factorial en forma matricial, dado anteriormente en la expresión (1) es:

$$x = \Lambda y + e$$

donde x es un vector columna de p variables, y es un vector columna de k factores, e es un vector de p residuos, que representa los efectos combinados de los factores específicos y los errores aleatorios de medición, finalmente $\Lambda=(\lambda_{ik})$ es una matriz de cargas factoriales de pxk.

Igualmente se señaló anteriormente que los residuos e no estan correlacionados entre sí y tampoco con los factores y, que las variables de los 3 vectores x, y y e están estandarizados con media cero y sus matrices de varianza-covarianza son denotadas respectivamente por $\Sigma_{(pxp)}$, $\Phi_{(kxk)}$ y $\Psi^2_{(pxp)}$. La matriz Ψ^2 es una matriz diagonal con elementos ψ^2_{ii}; i=1,2,. . .,p que son las varianzas únicas de las variables. Los factores tienen varianzas unitarias por lo que los elementos de la diagonal de Φ son todos iguales a uno. Además, si para k>1 los factores comunes son ortogonales, lo cual significa que no están correlacionados, entonces los elementos fuera de la diagonal de Φ son ceros y

Φ es una matriz identidad I. Por otro lado si los factores están correlacionados, Φ es la matriz de correlación de los factores. De la ecuación (1) que representa el modelo del análisis factorial y los supuestos expuestos anteriormente, la matriz de covarianza Σ de las variables manifiestas u observadas x, es la siguiente:

$$\Sigma = \Lambda\Phi\Lambda' + \Psi \qquad (13)$$

donde

$$\Phi_{(kxk)} = \begin{pmatrix} 1 & & & & \\ \phi_{21} & 1 & & & \\ \phi_{31} & \phi_{32} & 1 & & \\ \cdot & \cdot & \cdot & & \\ \cdot & \cdot & & \cdot & \\ \cdot & \cdot & & \cdot & \\ \phi_{k1} & \phi_{k2} & \cdots & \phi_{k,k-1} & 1 \end{pmatrix}$$

$$\Psi = \begin{bmatrix} \psi_1^2 & 0 & 0 \\ 0 & \psi_2^2 & 0 \\ 0 & 0 & \psi_p^2 \end{bmatrix} \qquad \Lambda_{(pxk)} = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1k} \\ \lambda_{21} & \lambda_{22} & \cdots & \lambda_{2k} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \lambda_{p1} & \lambda_{p2} & \cdots & \lambda_{pk} \end{bmatrix}$$

Las ecuaciones (1) y (13) representan un modelo para una población de individuos u observaciones. Esta población la caracterizan los parámetros Λ, Φ y Ψ . Pero como sabemos, en la práctica estos parámetros son desconocidos y podran estimarse a partir de una muestra de N individuos. Sea x_{ai} el valor observado de la variable i para un individuo a. Entonces los datos muestrales pueden escribirse como una matriz de observaciones **X** de orden Nxp. De la cual es posible calcular la media muestral del vector; $\bar{x}' = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_p)$ y la matriz de covarianza muestral $\mathbf{S}=(s_{ij})$, donde

$$\bar{x}_i = \left(\frac{1}{N}\right) \sum_{a=1}^{N} x_{ai} \qquad (14)$$

$$s_{ij} = \left(\frac{1}{n}\right) \sum_{a=1}^{n} \left(x_{ai} - \bar{x}_i\right)\left(x_{aj} - \bar{x}_j\right) \qquad (15)$$

con $n = N-1$

La información dada por la matriz S también puede representarse por una matriz de correlación $\mathbf{R} = (r_{ij})$ y un conjunto de desviaciones estándar $s_1, s_2, \ldots s_p$ donde $s_i = \sqrt{s_{ii}}$ y $r_{ij} = s_{ij} / \sqrt{s_i s_j}$

Generalmente en muchas aplicaciones, las unidades o escalas de medición de las variables manifiestas u observadas $\boldsymbol{x}$ son diferentes por lo cual se recomienda trabajar con la matriz de correlación **R**, por lo contrario cuando las variables $\boldsymbol{x}$ tienen todas la misma unidad o escala de medición se puede trabajar con la matriz de covarianza S.

Una vez puntualizados estos conceptos, retomemos la pregunta de estimación planteada al principio de este capítulo: ¿Cómo ajustar la matriz modelo Σ de la forma (13) a una matriz de covarianza muestral S?. En otras palabras debemos estimar Λ y Ψ en forma tal que la ecuación siguiente:

$$\Sigma = \Lambda\Phi\Lambda' + \Psi$$

sea tan cercana a la matriz S como sea posible. Primeramente suponemos que el número de factores k es conocido y en consecuencia puede especificarse "a priori". Sin embargo la mayoría de los estudios factoriales son de naturaleza exploratoria al menos en su primera etapa, y por lo tanto, el valor de k es desconocido.

En un análisis exploratorio el investigador o experimentador explora los datos empíricos para descubrir futuras características y relaciones interesantes sin la imposición de cualquier modelo o estructura definitiva de los datos.

Justamente en el análisis exploratorio puede surgir una estructura, o un modelo a partir del cual se pueden generar hipótesis. Pero por otro lado en un análisis confirmatorio, el investigador ha obtenido tal conocimiento sobre las variables manifiestas u observadas que lo colocan en la posición de formular hipótesis en donde se especifican las estructuras (variables latentes) subyacentes a las variables manifiestas y/o bien el investigador desea realizar un diseño confirmatorio contra una teoría específica o con estudios exploratorios previos. Sin embargo no obstante que en términos teóricos la diferencia es clara entre ambos, Jöreskog (1979, pp. 45-46) señala que en la práctica, está distinción no siempre es comprensible, pues muchos investigadores emplean ambos análisis debido a que tienen algunas variables conocidas y otras desconocidas, añadiendo además que sería muy deseable que las hipótesis sean producto de análisis exploratorios significativos de tal forma que estas puedan posteriormente ser rechazadas o no rechazadas, obteniendo información nueva, la cual deberá ser sujeta a técnicas estadísticas más rigurosas.

De cualquier forma el investigador quiere determinar el menor valor de k, es decir, minimizar k de tal forma que sea el mejor ajuste a sus datos (*principio de parsimonia*), el cual señala que para dos o más modelos igualmente compatibles para un grupo de datos, el modelo más simple será tomado como el mejor, es decir, en análisis factorial, sólo el modelo que tenga el menor número de factores será considerado el apropiado. El procedimiento común o usual es ir probando secuencialmente el incremento del valor de k hasta que se tenga un ajuste *"suficientemente bueno"* que Jöreskog (1975) cita es el método iterativo desarrollado por Lawley y Maxwell (1971).

Los métodos de estimación más comunes para el modelo de análisis factorial son los siguientes:

- Unweighted Least Squares (ULS) (Mínimos cuadrados no ponderados).
- Generalized Least Squares (GLS) (Mínimos cuadrados generalizados).
- Generally Weighted Least Squares (WLS) (Mínimos cuadrados ponderados generalizados).
- Maximum Likelihodd (ML) (Máxima verosímilitud).

Las estimaciones ULS, GLS, ML, WLS se obtienen de procedimientos iterativos que minimizan una función de ajuste particular para mejorar sucesivamente los parámetros estimados.

1) Método de mínimos cuadrados o también llamados mínimos cuadrados no ponderados. (ULS).

Minimiza la suma de cuadrados de todos los elementos de $(S - \Sigma)$:

$$U = tr(S - \Sigma)^2 = tr(S - \Sigma)'(S - \Sigma) \qquad (16)$$

donde:

tr denota la traza de una matriz.

y

$$S_{(pxp)} = \begin{pmatrix} s_{11}s_{12}\ldots s_{1p} \\ s_{21}s_{22}\ldots s_{2p} \\ . \\ . \\ . \\ s_{p1}s_{p2}\ldots s_{pp} \end{pmatrix} \qquad \Sigma_{(pxp)} = \begin{pmatrix} \sigma_{11}\sigma_{12}\ldots\sigma_{1p} \\ \sigma_{21}\sigma_{22}\ldots\sigma_{2p} \\ . \\ . \\ . \\ \sigma_{p1}\sigma_{p2}\ldots\sigma_{pp} \end{pmatrix}$$

Para obtener los estimadores de Λ, Ψ y Φ se emplean las siguientes ecuaciones siempre que $\Phi=I$:

$$\left(S-\hat{\Psi}\right)\hat{\Lambda} = \hat{\Lambda}\left(\hat{\Lambda}'\hat{\Lambda}\right) \tag{17}$$

y

$$\hat{\Psi} = diag(S-\hat{\Lambda}\hat{\Lambda}) \tag{18}$$

Estas dos ecuaciones sugieren el empleo de un esquema iterativo el cual se inicia con una estimación inicial de $\hat{\Psi}$: los k eigenvectores de S-$\hat{\psi}$ producen las columnas de $\hat{\Lambda}$ de la ecuación (17), y una estimación mejor se obtiene para $\hat{\psi}$ (donde $\hat{\Psi}$ es la matriz estimada de varianza covarianza del vector aleatorio de residuos) la cual se substituye en la expresión (18), continuándose este proceso iterativo hasta que se alcanza la convergencia. A este método se le conoce como *Factor Principal* ampliamente usado hasta el surgimiento de las computadoras de alta velocidad en la década de los años 70's. Si $\Psi=0$ el método se reduce a un sólo paso, al análisis de componentes principales y es por ello que en algunos paquetes estadísticos, como el SPSS se ofrece como una opción en el análisis factorial, pero sólo en este caso.

2).-Método de Mínimos cuadrados generalizados (GLS).

Este método puede considerarse como una generalización del método de estimación de mínimos cuadrados en donde se busca minimizar la suma de cuadrados de todos los elementos de $([S-\Sigma]S^{-1})$.

$$G = tr([S-\Sigma]S^{-1})^2 \tag{19}$$

donde *tr* es la traza de una matriz.

De manera análoga este método se dirige a la estimación de las mismas ecuaciones anteriores (17) pero remplaza la ecuación (18) por la siguiente expresión:

$$\mathrm{diag}\{S^{-1}(\hat{\Lambda}\hat{\Lambda}' + \hat{\Psi}) - S]S^{-1}\} = \mathrm{diag}\{\ 0\ \} \qquad (20)$$

con un proceso iterativo hasta alcanzar la convergencia.

3).-Método de Mínimos Cuadrados Ponderados (WLS).

Es una generalización del método de mínimos cuadrados.

$$WLS = tr([S - \Sigma]\Psi^{-1})^2 \qquad (21)$$

Se le llama método de mínimos cuadrados ponderados ya que pondera la diferencia de $S - \Sigma$ por los valores de Ψ^{-1}. La estructura de esta ecuación (21) es la misma que las ecuaciones de estimación (16) y (17), sólo que $\hat{\Lambda}$ es reemplazada por $\hat{\Psi}^{-1/2}\hat{\Lambda}$ y con S también reemplazada por $\hat{\Psi}^{-1/2}S\hat{\Psi}^{-1/2}$.

4).-Método de Máxima Verosimilitud (ML).

La ecuación de máxima verosimilitud queda expresada como sigue: (Krzanowski y Marriott 1995)

$$L=(2\pi)^{-np/2}\,|\Lambda\Lambda' + \Psi|^{-n/2} \exp -1/2\sum_{i=1}^{n}(x_i - \mu_i)'(\Lambda\Lambda' + \Psi)^{-1}(x_i - \mu_i) \qquad (22)$$

Maximizando esta función respecto a μ se obtiene que el estimador máximo verosímil es $\hat{\mu} = \bar{x}$. Los demás parámetros se pueden estimar independientemente de la media y se obtienen al maximizar la función de verosimilitud (22) equivalente a maximizar el logaritmo de la función :

$$L' = \frac{-1}{2} np \log(2\pi) - \frac{1}{2} n \log|\Lambda\Lambda' + \Psi| - \frac{1}{2} n\, tr[(\Lambda\Lambda' + \Psi)^{-1} S)] \qquad (23)$$

donde:

$$S = \frac{1}{n}\sum_{i=1}^{n}[(x_i - \bar{x})(x_i - \bar{x})'] \qquad (24)$$

que es la matriz de varianza covarianza de los datos observados

Aquí se requiere maximizar a L' respecto a Λ y a Ψ. Al maximizar se llega al conjunto de ecuaciones simultáneas siguiente:

$$(\hat{\Psi}^{-1/2} S \hat{\Psi}^{-1/2})(\hat{\Psi}^{-1/2} \hat{\Lambda}) = (\hat{\Psi}^{-1/2} \hat{\Lambda})(I + \hat{\Lambda}' \hat{\Psi}^{-1} \hat{\Lambda}) \qquad (25)$$

$$\hat{\Psi} = \mathrm{diag}(\hat{\Lambda}\hat{\Lambda}' - S) \qquad (26)$$

Estas ecuaciones no tienen solución analítica o explícita y se requiere de un método numérico iterativo de solución.

Una característica importante de este método es que es invariante ante la escala de medición de las variables manifiestas u observadas. Se trabaja con las variables en la escala original, equivalentemente con la matriz de varianza covarianza S, y se obtienen estimadores $\hat{\lambda}_{ij}$, $\hat{\Psi}_i$, entonces $\hat{\lambda}_{ij}/\sqrt{s_{ii}}$, $\hat{\Psi}_i / \sqrt{s_{ii}}$ son los estimadores correspondientes que se obtienen al trabajar con la matriz de correlación **R**. Los métodos GLS y ML tienen la propiedad de ser invariantes a la escala o unidad de medición no así para el método ULS.

Así como en el método factor principal no se requieren supuestos sobre la distribución, en el método de máxima verosimilitud sí los hay; se asume que el vector p dimensional de variables manifiestas x tiene una distribución normal p dimensional con vector media $\bar{x} = 0$ y matriz de varianza covarianza $\Sigma = \Lambda\Lambda' + \Psi$. Sin embargo se ha encontrado que los estimadores son robustos ante la falta de normalidad (ver los estudios de Monte Carlo realizados por Boomsma (1983) y Harlow (1985) y los estudios teóricos de Browne (1987) y Anderson y Amemiya (1985, 1986) citado en el manual de operación del paquete LISREL 7 (1989). La matriz de cargas factoriales Λ tiene k factores comunes, donde el valor de k se ha especificado antes de extraer los factores estimados.

La idea que subyace al estimador de máxima verosimilitud es la siguiente: se supone que se conoce la forma general de la distribución de la población de donde se extrajo la muestra, y esta es una distribución normal multivariada. Pero no conocemos los parámetros poblacionales que dan a esta distribución una forma particular entre todas las posibles distribuciones multivariadas normales. En ausencia de tal conocimiento, sin embargo, podemos tomar valores arbitrariamente y tratarlos como si ellos fueran los parámetros poblacionales y preguntamos qué es la *verosimilitud...* de observar ciertos valores para las variables en una sola observación extraída de dicha población. Si tenemos más de una observación, entonces podemos preguntamos ¿qué es la verosimilitud conjunta de dicha obtención de la muestra de vectores observados o manifiestos?. Finalmente podemos preguntar: ¿qué valores de los parámetros poblacionales hacen que las observaciones muestrales tengan la mayor verosimilitud conjunta?. Cuando contestemos a esta pregunta, tomaremos tales valores como *estimadores máximo verosímiles* de los parámetros poblacionales.

Jöreskog (1979, p. 19) señala que cada función ajustada ULS, GLS y ML se minimiza con respecto a Λ,Φ y Ψ. Dichas funciones pueden minimizarse numéricamente, básicamente por el mismo algoritmo. Cuando x tiene una distribución normal multivariada, los métodos GLS y ML producen estimaciones con buenas propiedades en muestras grandes, ambos métodos requieren de una matriz S positiva definida, mientras que el método ULS puede operar con una matriz semi-positiva.

Las funciones de ajuste antes mencionadas son siempre no-negativas, son iguales a cero sólo cuando hay un ajuste perfecto en cuyo caso la matriz $\hat{\Sigma}$ es igual a S.

Actualmente, dado el extraordinario avance de las computadoras, el método ML es ampliamente empleado en el análisis factorial. Cuando dicho método se emplea, puede iniciarse con un valor de k=1, secuencialmente se añaden valores sucesivos y se prueba la significancia del empleo de este valor en la siguiente expresión (27), empleando la prueba de razón verosimil como prueba de bondad de ajuste del modelo propuesto de k factores. La estadística de prueba es la siguiente (Kraznowski y Marriott cap. 12 ,1995):

$$w = n[tr(\hat{\Lambda}\hat{\Lambda}' + \hat{\Psi})^{-1}S - \log|(\hat{\Lambda}\hat{\Lambda}' + \hat{\Psi})^{-1}S| - p] \qquad (27)$$

Si $\Psi > 0$, w tiene una distribución χ^2 con grados de libertad igual a $\frac{1}{2}[(p-k)^2 - (p+k)]$

bajo la hipótesis nula siguiente:

$$Ho: \Sigma = \Lambda\Lambda' + \Psi \text{ vs. } Ha: \Sigma \neq \Lambda\Lambda' + \Psi, \text{ es decir, } \Sigma \text{ no está restringida}$$

donde:

$\Sigma = \Lambda\Lambda' + \Psi$ es la matriz de dispersión bajo el modelo factorial y S es la matriz común de dispersión

Krzanowski y Marriott (1995) citan que Bartlett (1954. JRSS-B, vol 16, pp 296-298) mostró que la expresión (27) se puede aproximar asintóticamente a la distribución Ji-cuadrada si en ella se reemplaza a n por la siguiente expresión:

$$n - 1 - \frac{1}{6}(2p + 5) - \frac{2}{3}k \qquad (28)$$

Cuando el resultado del valor de Ji-cuadrada no es significativo se termina el proceso concluyéndose que el ajuste al modelo es adecuado. Es posible que teniendo una solución *"prohibitiva"* con $(p-k)^2 > p+k$ proporcione un ajuste adecuado.

SELECCION DEL MODELO

En esta exposición hemos supuesto que el número de factores k en el modelo es conocido y en consecuencia se fija, pero realmente en muchas aplicaciones prácticas, este no es el caso: k será otro parámetro del modelo estimado para los datos. Antes de emplear el método de máxima verosimilitud en el análisis factorial debemos seleccionar un criterio que nos permita determinar el número de componentes significativos para el caso de análisis de componentes principales, contándose para ello con diversos criterios: a) los eigenvalores de la matriz de correlación **R**; b) porcentaje de explicación de la varianza; c) número de componentes con eigenvalores mayores a uno y d) componentes a partir de la gráfica de *"desmoronamiento" (scree plot)* explicada en la primera parte de este trabajo.

Si se emplea el método de máxima verosimilitud en análisis factorial, se usa la prueba de la razón máximo verosímil como anteriormente se señaló. Desde luego, es posible que teniendo una solución "prohibida" con $(p - k)^2 > p + k$ nos proporcione evidencia para no rechazar dicho modelo. También, aunque este método se considera objetivamente *"adecuado"* para determinar el valor de k, se presentan algunos problemas: no es estrictamente válido para las pruebas de hipótesis, ya que el nivel de significancia establecido se ve afectado por la naturaleza del proceso secuencial y por otro lado, el que la prueba no haya resultado significativa para un nivel de significancia establecido no significa que se haya encontrado el valor *"óptimo"* de k. El *"mejor"* modelo será aquel para el cual el valor de k, los en donde los datos ajustan *"mejor"* a los parámetros del modelo.

Determinar el *"mejor"* modelo involucra un intercambio entre el número de parámetros y la bondad del ajuste, Krzanowski y Marriott (1995) citan a Akaike (1983) el cual sugirió una alternativa para seleccionar k.

Por otro lado, cuando se emplea el método de máxima verosimilitud y se asocia con una prueba de significancia, el mayor tamaño de muestra posible empleado, dará una mejor aproximación a χ^2. Krznowski y Marriott (1995) señalan que Lawley y Maxwell(1971) sugirieron que la prueba es apropiada si la muestra contiene al menos 51 casos más, que el número de variables bajo consideración, esto es, n-p-1≥50, donde n es el tamaño de la muestra y p es el número de variables manifiestas u observadas, desde luego esta es sólo una regla empírica. Sin embargo cuando el tamaño de muestra es *"grande"*, el método de máxima verosimilitud y el método de factor principal convergen al mismo resultado.

Algunas ventajas del método de máxima verosimilitud sobre el método del factor principal son las siguientes: 1) en el método de ML los factores estimados son independientes de la escala de medición, de tal forma que no limita al investigador al uso de instrumentos con iguales escalas, y 2) en el método de ML las pruebas de significancia estadística para el número de factores puede derivarse empleando la χ^2 tal y como se mencionó anteriormente.

$\Phi = I$ y $\Lambda'\Lambda$ será diagonal en el método ULS.

$\Phi = I$ y $\Lambda'\Psi^{-1}\Lambda$ será diagonal en los métodos GLS y ML.

Esta libertad de elegir nos lleva a tener un conjunto arbitrario de factores que pueden rotarse, a otro nuevo conjunto de factores, lo que permite hacer una interpretación más fácil.

En estudios comparativos realizados entre diferentes métodos de estimación por Jöreskog (1963 pp. 129 y 130), señala que el método de máxima verosimilitud es el mejor de los métodos de estimación existentes. La alta eficiencia de este método está basada en el soporte computacional que implica. El método es iterativo, comenzando con una matriz inicial de cargas factoriales de forma tal que a partir de aproximaciones el proceso realiza sucesivas iteraciones hasta que converge en una matriz final. Con frecuencia la matriz inicial puede ser muy cercana o semejante a la matriz final resultando el proceso altamente satisfactorio. Sin embargo una desventaja es que pueden encontrarse números imaginarios en este proceso iterativo, puede converger muy lentamente y la estimación para algún Ψ_i puede tender a cero. Aunque se han hecho algunas mejoras, la estimación por Máxima Verosimilitud permanece impracticable para aquellos problemas con muestras "*pequeñas*". A pesar de esto, proporciona un método rápido y confiable en análisis factorial, por lo cual está instrumentado en la gran mayoría de los paquetes estadísticos, tales como SPSS y SAS.

Dos aspectos involucrados en este procedimiento anterior requieren de una breve explicación:

1) la elección de los valores iniciales para $\hat{\psi}_i$, $i=1,\ldots,p$, para comenzar el proceso iterativo, y

2) la tendencia para algunos $\hat{\psi}_i$, a cero en aplicaciones prácticas (este último mencionado anteriormente).

Considerando el primer aspecto, si consideramos el modelo $\boldsymbol{x} = \mu + \Lambda \boldsymbol{y} + \boldsymbol{e}$ como expresando cada x_i en términos de una parte "sistemática" (factores comunes y_i) y un *error* e_i, entonces podemos considerar cada x_i como una función lineal de todas las otras x's más un residuo. La comunalidad de la i-ésima variable, $h^2_i = \mathrm{var}(x_{ii}) - \psi_y$, no puede ser menor que $\mathbf{R}^2_i\,\mathrm{var}(x_i)$ (donde R^2_i es el coeficiente de correlación múltiple entre x_i y todas las otras x's), y las dos con frecuencia cercanas al valor. Así un valor inicial citado en Krzanowski y Marriott (1995) para ψ_i sería el siguiente: $\mathrm{var}(x_i)\{1 - \mathbf{R}^2_i\}$, $i=1,\ldots,p$. Notése, sin embargo, que existen otros valores iniciales simples por ejemplo, Jöreskog (1963) sugiere usar $(1 - \frac{1}{2}k/p)(1/s^{ii})$ donde s^{ii} es el i-ésimo elemento de la diagonal $\mathbf{S}^{-1}$.

En cuanto al segundo aspecto, la aproximación simple de igualar las derivadas parciales a cero y resolviendo para los parámetros desconocidos nos puede llevar (sorpresivamente a menudo) a una solución con uno o más ψ_i negativos (solución inapropiada, también llamada *casos Heywood*").

Numéricamente, esta situación se aborda, forzando la optimización del logaritmo de la función de máxima verosimilitud sobre $\psi_i \geq \varepsilon$ para algún valor predeterminado $\varepsilon > 0$ y entonces posiblemente volviendo a trabajar el análisis con $\psi_i = 0$ para explorar las soluciones limitadas. Krzanowski y Marriott (1995) citan que Van Driel (1978) sugiere tres posibles razones explicativas:

1) debido a variación muestral;

2) porque ningún modelo factorial interpretable ajusta a las observaciones o

3) porque muchas cargas factoriales verdaderas son cero.

La solución limitada es apropiada para el primer caso, pero en las otras dos situaciones una selección inadecuada de variables para el modelo pudiese haber ocurrido, las variables "transgresoras" (esto es, para las que $\psi_i \to 0$) pueden eliminarse del análisis, o bien pueden incluirse nuevas variables o factores.

Antes de dejar este tema sobre las estimaciones mencionaremos brevemente y de manera muy general dos soluciones del análisis factorial que no son propiamente métodos de estimación como veremos a continuación y que son opciones presentadas en algunos paquetes estadísticos tal como el SPSS.

1).-ALPHA

En los casos de la estimación de mínimos cuadrados y el de máxima verosimilitud, se asumió que las variables manifiestas consideradas por el investigador, constituyen el universo y que el muestreo es sólo de los sujetos o individuos. En el método ALPHA, sin embargo, las variables incluídas en el análisis factorial son consideradas una muestra del universo de *variables*, mientras se asume que estas variables son observadas de una *población* dada de individuos. Por lo tanto, en el método ALPHA, el énfasis es en la inferencia psicométrica y no en la inferencia estadística en el sentido usual del que se hace y que aquí hemos hecho referencia.

Este método produce primeramente una matriz de correlación ajustada o corregida por *"atenuación"*:

$$R_1 = H^{-1} (R - U^2) H^{-1} \qquad (31)$$

donde:

R = matriz de correlación.

R_1 = matriz de correlación corregida o ajustada, para la solución ALPHA

U^2 = matriz diagonal de componentes únicos (*unique components)*

H^{-1}= matriz diagonal cuyos elementos son los recíprocos de las raíces cuadradas de las comunalidades.

H^{-1}= matriz diagonal cuyos elementos son los recíprocos de las raíces cuadradas de las comunalidades.

Entonces la ecuación de resolución asociada con esta matriz ajustada es como sigue:

$$det(R_1 - I) = 0$$

donde *det* es el determinante.

En este método ALPHA la solución se encuentra a través de la comunalidad, dando mayor peso a las variables con mayor comunalidad, la solución es entonces complicada por el hecho de que se parte de una comunalidad e itera con estas variables para llegar a la solución final.

En este método el número de factores que se retienen está determinado por el criterio de que los eigenvalores asociados sean mayores a 1. Este criterio es equivalente al criterio de que el *coeficiente de generalización asociado* α (de ahí el nombre de ALPHA) en el universo de variables debera ser mayore a cero. Aquí, desde luego, la prueba de significancia no es en el sentido usual, ya que se asumió que los individuos son la población.

2.-IMAGEN

El análisis de la solución IMAGEN se fundamenta en la distinción entre la parte común de una variable y la parte única. La parte común de una variable es definida como la parte que es predecible por una combinación lineal de todas las otras variables en el grupo, y es llamada la IMAGEN de la variable. La parte única es la parte de la variable que no es predecible por la combinación lineal de otras variables a la que se le llama ANTI-IMAGEN.

El análisis de la IMAGEN también supone que este universo de variables es potencialmente infinito. Las variables manifiestas u observadas pueden considerarse como una muestra tomada de una población infinita de variables. Sin embargo, si tomamos todas las variables de un universo potencial para su examen, el cuadrado de la imagen de una variable será equivalente a la *comunalidad* de una variable definida en el análisis factorial, y el *cuadrado* de la ANTI-IMAGEN de una variable será equivalente a la *varianza única* (asumiendo que las variables están estandarizadas). En otras palabras, la correlación múltiple al cuadrado entre una variable y el resto de las variables en este universo es el mismo como la comunalidad de una variable dada.

La IMAGEN y ANTI-IMAGEN definidas por una muestra de variables son llamadas imágenes y anti-imágenes parciales respectivamente. Aunque las primeras sólo aproximan la imagen total, está completamente especificada por las variables manifiestas u observadas. En este sentido, en el análisis factorial la parte común de una variable es definida como una combinación lineal de factores hipotéticos, y nunca como una función exacta de las propias variables observadas.

Dada una muestra de variables y sus correlaciones, el análisis IMAGEN construye una matriz de covarianza de imagen parcial, que es dado por:

$$R_2 = (R - Z^2)R^{-1}(R - Z^2) \tag{32}$$

donde

R_2 = Matriz de correlación corregida, para la solución IMAGEN
R = Matriz de correlación.
Z^2 = Matriz diagonal, cuyos elementos constituyen la varianza de cada variable no explicada por las otras variables, o la varianza ANTI-IMAGEN.

El proceso involucrado en la ecuación (32) es el siguiente:

1) reemplazar en al diagonal principal de R la correlación múltiple al cuadrado de cada variable con cada una de las restantes y

2) reajustar los elementos fuera de la diagonal, en el sentido de hacer que el resultado sea una matriz positiva definida. Entonces la *eigenecuación* es aplicada a la matriz siguiente:

$$det\,(R_2 - I) = 0 \tag{33}$$

El número de factores a extraer o retener no esta dado por el examen de los eigenvalores de la ecuación (33). Usualmente, el número de factores extraídos de esta forma es relativamente grande, generalmente toma la mitad de las variables en el análisis. Kaiser sugirió que los factores insignificantes o no interpretables deberán ser eliminados después de las rotaciones.

Finalmente diremos que si una proporción "grande" de coeficientes tienen una alta correlación en la matriz ANTI-IMAGEN deberá reconsiderarse el uso del análisis factorial.

CAPITULO 4

ROTACION FACTORIAL

Anteriormente se mencionó el concepto de la no determinación del modelo factorial diciendo que este era un problema que podría resolverse mediante la estimación del *espacio factorial* eligiendo alguno de los métodos diseñados para efectuar una *rotación*.

El término *"rotación"* en análisis factorial significa lo mismo que *"rotar"*, ¿y qué es rotar?, rotar significa que teniendo como ejes de referencia a los factores estos últimos son girados de su origen hasta alcanzar alguna cierta posición. El caso más conocido y sencillo es el de la rotación ortogonal , en la cual los ejes se mantienen a un ángulo de 90°, pero también puede suceder que se roten los ejes sin conservar el ángulo de 90° conocida como rotación oblicua.

¿Cuál es el principal objetivo de rotar los factores?.

Alcanzar soluciones factoriales significativas (es decir, interpretables), en pocas palabras, simplificar la estructura factorial. Cuando lo que se busca es solamente reducir variables, los factores no se rotan y en el caso particular de las matrices factoriales no rotadas, simplemente se está interesado en la mejor combinación lineal de las variables manifiestas u observadas.

Independientemente del tipo de rotación que se elija el "patrón" más definido, será aquel que dé estructuras de cargas estimadas $\hat{\lambda}_{ij}$ más sencillas que proporcionen una mejor comprensión del análisis que se haga. Thurstone (1947) consideró varios criterios que este patrón deberá tener. Actualmente se dispone de diferentes rotaciones en el uso de paquetería sobre análisis factorial, que cumplen con estos criterios. La forma como estas rotaciones operan es mediante la especificación de un índice de "simplicidad" en términos de las cargas factoriales de las variables manifiestas y entonces buscar la transformación que optimice este índice.

Las 2 rotaciones ortogonales más usadas son:
1) **VARIMAX** (Kaiser).
2) **QUARTIMAX** (Neuhaus y Wrigley)

La primera rotación VARIMAX busca simplificar las columnas en la matriz factorial, y la simplificación máxima se consigue cuando hay sólo unos o ceros en una sola columna. En este método los factores se rotan de forma tal que se maximiza la suma de las varianzas de las cargas factoriales al cuadrado con cada columna de la matriz de cargas factoriales rotadas.

Una de las formas de buscar transformaciones ortogonales es el método VARIMAX. El índice de simplicidad es el siguiente:

$$\mathbf{R}_1 = \frac{1}{p^2} \sum_{j=1}^{k} \{k\sum_{i=1}^{p} b^4_{ij} - (\sum_{i=1}^{p} b^2_{ij})^2\} \qquad (34)$$

donde $b_{ij} = \lambda_{ij}/(\sum_{j=1}^{k} \lambda^2_{ij})^{1/2}$

y λ_{ij} = cargas factoriales.

Por lo que respecta a la segunda rotación ortogonal, es decir, QUARTIMAX su objetivo es simplificar los renglones de una matriz factorial, el objetivo en este tipo de rotación es conseguir que las cargas factoriales de las variables manifiestas u observadas sean lo más alto posible en un sólo factor y lo más bajas posibles en los otros factores, es decir, que carguen alto o muy cercanas a lo alto en el mismo factor lo cual nos llevará a la obtención de un sólo factor llamado "*factor general*", lo cual ocasiona la inevitable relación entre muchas variables.

La transformación se busca maximizando el índice siguiente:

$$\mathbf{R}_2 = \sum_{i=1}^{p} \sum_{j=1}^{k} \lambda^4_{ij} - \frac{1}{pk}(\sum_{i=1}^{p} \sum_{j=1}^{k} \lambda^2_{ij})^2 \qquad (35)$$

En una matriz factorial, las columnas representan factores (k) con cada renglón correspondiente a una variable manifiesta (p) con su respectiva carga factorial en el factor. Por simplificación de renglones, se entiende que muchas variables en cada renglón tengan cargas factoriales tan cercanas a cero como sea posible. Por simplificación de columnas lo que se busca es que muchas cargas factoriales de las variables en cada columna sean tan cercanas a cero como sea posible.

La rotación VARIMAX tiende a ser más estable que la que se obtiene por el método QUARTIMAX.

Pero, ¿qué sucede si hay razones teóricas y/o empíricas que permiten suponer que los factores son correlacionados?, la rotación ortogonal no opera en estas situaciones, teniendo que recurrir a otro tipo de transformaciones tales como la rotación oblicua.

Aunque se mencionó anteriormente algunas de las ventajas de la rotación oblicua sobre la ortogonal, presenta sin embargo ciertos problemas: dado que las soluciones oblicuas se obtienen a partir de las correlaciones entre factores, su interpretación es más compleja, particularmente porque se tiene que asumir la presencia de factores causales de orden superior para explicar las correlaciones entre los factores. Existen varios métodos de aproximaciones a la rotación oblicua, veremos a continuación dos de los más comunes:

1) Método que emplea ejes de referencia

2) Método que emplea la matriz patrón.

Con respecto al primero de ellos el cual fue desarrollado por Thurstone en el año de 1947 es un método en donde todas las soluciones incluidas están basadas en los siguientes supuestos: 1) sí hay grupos definidos de variables que representan dimensiones separadas, y 2) sí estos grupos están correctamente identificados por factores primarios, entonces cada grupo de variables tendrá proyecciones cercanas a cero en todos los ejes menos en uno, teniendo en consecuencia un criterio llamado QUARTIMIN.

$$\mathbf{N} = \sum_{i=1}^{p} \sum_{j<l=1}^{k} a_{ij}^2 a_{il}^2 \qquad (36)$$

donde a_{ij} y a_{il} son proyecciones en el j-ésimo y l-ésimo ejes de referencia.

Este valor será igual a cero si todas las variables son unifactoriales. Lo que buscamos en la rotación oblicua son cargas factoriales que minimicen **N** (en la rotación ortogonal, este criterio es equivalente a QUARTIMAX). Paralelo a la modificación VARIMAX del criterio QUARTIMAX en rotaciones ortogonales, hay criterios conocidos como COVARIMIN o BIQUARTIMIN. El valor que se busca minimizar es la covarianza de los elementos al cuadrado de las proyecciones en los ejes.

$$\mathbf{C} = \sum_{j<l=1}^{k} \left(p\sum_{i=1}^{p} a^2_{ij} a^2_{il} - \sum_{i=1}^{p} a^2_{ij} \sum_{i=1}^{p} a^2_{il}\right) \qquad (37)$$

Se normalizan si reemplazamos a^2_{ij} con a^2_{ij}/h^2_i , donde $h^2_i = \sum^{k} \lambda^2_{ij}$ y k es el número de columnas en la matriz patrón y λ_{ij} es la carga factorial de la variable i en el factor j de la matriz patrón. Cuando se aplica al mismo conjunto de datos, el criterio COVARIMIN produce pocos factores oblicuos, mientras que el criterio QUARTIMIN dá más factores oblicuos.

En cuanto al segundo método de rotación oblicua, el método que emplea la matriz patrón fue desarrollado por Jennrich y Sampson , básicamente lo que hace es simplificar las cargas en los factores primarios. Antes de continuar con esta exposición es necesario definir algunos conceptos relacionados con la rotación oblicua, los cuales son presentados a continuación.

En el método de rotación oblicua se busca el mínimo valor definido en la siguiente ecuación:

$$\mathbf{D} = \sum_{j=l=1}^{k} \left[\sum_{i=1}^{p} \lambda^2_{ij} \lambda^2_{il} - \mathrm{d}\left(\sum_{i=1}^{p} \lambda^2_{ij} \sum_{j=1}^{p} \lambda^2_{jl}\right)/n\right] \qquad (38)$$

En general la relación entre la *matriz patrón* (*pattern matrix*) y la *matriz estructura* (*structure matrix*) (Jöreskog 1988, p. 92) es :

$$\hat{\Lambda} = \textit{matriz patrón} \quad \text{y} \quad \hat{\Lambda}\hat{\Phi} = \textit{matriz estructura}$$

En general en esta rotación oblicua, valores grandes de **D** en la expresión (38) dan más soluciones de esta naturaleza y valores (negativos) menores producen mayor número de soluciones ortogonales. Si el patrón factorial es unifactorial (lo más simple posible), la especificación de d=0 identifica un patrón correcto.

En conclusión las cargas de los elementos de la matriz patrón reflejan sólo la contribución directa de un factor a una variable y no la contribución total que es indexada por la matriz estructura. Con altas correlaciones factoriales, es totalmente posible que una variable sea "*explicada*" por un factor, ya que puede tener una alta carga λ en un sólo factor y baja en los otros factores. Sin embargo, estas cargas λ's pueden en ocasiones ser altas en todos los factores.

Generalmente la mayoría de los experimentadores o analistas interpretan los elementos de la matriz patrón más que los elementos de la matriz estructura. Y entre los 2 métodos, el más empleado es justamente el de la matríz patrón, ya que el de la matriz estructura implica efectuar la rotación manual, puesto que los ejes son determinados por el propio analista. Además es importante tener presente que los elementos de la matriz patrón son pesos o ponderaciones en la predicción de la variable observada x_i del factor y_j. En tanto que los elementos de la matriz estructura son correlaciones entre la variable x_i y el factor y_j. En el caso de la rotación ortogonal los valores de ambas matrices son equivalentes, no así para la rotación oblicua. Desde el punto de vista práctico, las ponderaciones de la matriz patrón pueden exceder a uno puesto que no son correlaciones, en tanto que las cargas de la matriz estructura si lo son, por lo tanto no exceden a uno y son fácilmente interpretables.

TRANSVARIMAX.

Mulaik (1972) cita que Saunders (1962), llevó a cabo una serie de trabajos tendientes a integrar los métodos **QUARTIMAX** y **VARIMAX** en un sólo programa computacional que incluye las rotaciones **OBLIMAX OBLICUA**, descubriendo que dadas las restricciones de la rotación ortogonal en cuanto a su representación en un sólo plano, la tangente de un ángulo deseable de rotación puede ser la raíz cuadrada de la siguiente ecuación cuártica:

$$x^4 + \beta x^3 - 6x^2 - \beta x + 1 = 0 \tag{39}$$

$$\text{donde } \beta = \frac{\left(\sum_{i=1}^{n} a_{ij}^4 - 6\sum_{i=1}^{n} a_{ij}^2 a_{ik}^2 + \sum_{i=1}^{n} a_{ik}^4\right) + K4\left(\sum_{i=1}^{n} a_{ij}a_{ik}\right)^2 - \left(\sum_{i=1}^{n} a_{ik}^2 - \sum_{i=1}^{n} a_{ij}^2\right)^2]/n}{\left(\sum_{i=1}^{n} a_{ij}^3 a_{ik} - \sum_{i=1}^{n} a_{ij}a_{ik}^3\right) + K\left[\sum_{i=1}^{n} a_{ij}a_{ik}\left(\sum_{i=1}^{n} a_{ik}^2 - \sum_{i=1}^{n} a_{ij}^2\right)\right]/n} \tag{40}$$

donde:

K es una constante que en la rotación varimax es igual a uno, mientras que en el método QUARTIMAX es igual a cero; a_{ij} y a_{ik} son elementos de la j y k ésima columna de la matriz estructura normalizada. Esta matriz empleada por el método de rotación VARIMAX es $\mathbf{H}^{-1}\mathbf{A}$ donde $\mathbf{H}^2$ es una matriz de comunalidades (nxn) y $\mathbf{A}$ es la matriz estructura inicial nxr no rotada. Donde n es el número de variables y r es el número total de factores.

El coeficiente β en esta ecuación variará de acuerdo al método de rotación ortogonal involucrado. Asimismo Saunders se ocupó de estudiar los efectos cuando se usaban valores para K>1. Encontrando que el valor óptimo de K quedaba definido por la siguiente expresión:

$$K = \frac{\sum_{j=1}^{r} \sum_{i=1}^{n} a_{ij}^4}{\sum_{j=1}^{r} \left(\sum_{i=1}^{n} a_{ij}^2\right)^2} \tag{41}$$

que es la razón de los mismos términos que para el criterio VARIMAX de los cuales Saunders toma la diferencia y lo llama criterio RATIOMAX.

Este valor de K fue determinado por los datos después de cada ciclo de rotaciones en un sólo plano con lo cual fue posible llegar a una solución RATIOMAX final y observar su proceso de convergencia.

EQUAMAX

La experiencia de Saunders con la definición de K en el criterio RATIOMAX mostró que desafortunadamente, este no producía resultados satisfactorios. Frecuentemente en las iteraciones K puede alcanzar un máximo rápidamente y entonces decrecer en iteraciones subsecuentes, particularmente en los casos cuando se creyó que también muchos factores fueron aceptados por la rotación.

Sin embargo, Saunders observó que el método RATIOMAX producía los resultados más satisfactorios, cuando alcanzaba un valor de K igual a ½ del número de factores rotados, esto es, cuando:

$$K = r/2 \tag{42}$$

Saunders decidió llamar a esta solución EQUAMAX.

PARSIMAX

Más exactamente Mulaik (1986) señala que Crawford (1967) buscó combinar la simplificación-renglón característica de la rotación VARIMAX en un método de rotación ORTOGONAL conocida como el método PARSIMAX (esto es, máxima parsimonia). El método PARSIMAX produce resultados que para algunos datos son muy similares a la solución EQUAMAX y así arrojar alguna información sobre la razón o causa de por qué EQUAMAX produce soluciones de estructura simple.

Crawford considera que un índice de simplificación en los renglones es el criterio q_1 de Carroll, esto es, $q_1 = \sum^{n}_{i=1} \sum^{r}_{j=1} \sum^{r}_{k=1} a_{ij}^2 a_{ik}^2$, $j \neq k$. Cuando q_1 es un mínimo cada variable es explicada por pocos factores comunes; por lo tanto se tiene la máxima parsimonia usando los factores que explican la varianza común de las variables.

Los métodos de rotación disponibles para el paquete SPSS son: VARIMAX, EQUAMAX, QUARTIMAX y OBLIMIN y para el SAS son: VARIMAX, EQUAMAX, ORTHOMAX, QUARTIMAX y PARSIMAX.

Con el fin de ilustrar las rotaciones factoriales ortogonal y oblicua se presenta la reproducción de un ejemplo hipotético desarrollado por Bartholomew (1985) . En este ejemplo se observa que la rotación oblicua recupera "*mejor*" o se "*aproxima más*" al modelo o estructura originalmente planteado en la matriz de correlación que la rotación ortogonal. Para ello se partió de la siguiente matriz de correlación de seis variables manifiestas: x_1, x_2, x_3,x_4,x_5, x_6 .

MATRIZ DE CORRELACION

	x_1	x_2	x_3	x_4	x_5	x_6
x_1	1.0000					
x_2	.8636	1.0000				
x_3	.6792	.6738	1.0000			
x_4	.5079	.5164	.3274	1.0000		
x_5	.6347	.5738	.4165	.6751	1.0000	
x_6	.6396	.6232	.4614	.7377	.8384	1.0000

Tabla No. 7

La matriz de correlación se calculó a partir de una matriz de datos de orden 100x6. Las observaciones originales fueron generadas a patir de 100 observaciones

El modelo *Gaussiano* propuesto tiene la estructura bidimensional (k_1, k_2) con $\rho(k_1, k_2) = .707$ siguiente:

$$
\begin{aligned}
x_1 &= 0.949\,k_1 + e_1 \\
x_2 &= 0.894\,k_1 + e_2 \\
x_3 &= 0.707\,k_1 + e_3 \\
x_4 &= 0.707\,k_2 + e_4 \\
x_5 &= 0.894\,k_2 + e_5 \\
x_6 &= 0.949\,k_2 + e_6
\end{aligned}
$$

donde: $x_1, x_2, x_3, x_4, x_5, x_6$ son las variables manifiestas
$e_1, e_2, \ldots, e_6$ son los errores generados a partir de la normal estandarizada.

Los datos de esta matriz se procesaron con el paquete SPSS con rotación ortogonal (VARIMAX) y por la oblicua (OBLIMIN). Antes de presentar los datos obtenidos es preciso señalar que estos difieren ligeramente de los reportados por Bartholomew, debido a que en este trabajo se procesaron a partir de la matriz de correlación reportada y no a partir de la matriz de observaciones originales

Resultados obtenidos por el SPSS para la rotación ortogonal (VARIMAX) y la oblicua (OBLIMIN).

FACTORES NO ROTADOS

FACTOR MATRIX

VARIABLES	F1 λ_{i1}	F2 λ_{i2}
x_1	.87033	.34863
x_2	.84723	.36230
x_3	.64595	.34142
x_4	.71389	-.29666
x_5	.82791	-.30008
x_6	.88357	-.35553

Tabla No. 8

En esta tabla No. 8 se presenta la solución inicial sin la rotación de los factores, observándose que el Factor 1 es un factor *general*, en tanto que el segundo factor es *bipolar*, ya que tiene cargas factoriales positivas y negativos mayores a .30 consideradas significativas (este valor de .30 es una regla empírica particularmente en Psicología ver Kerlinger 1981).

ESTADISTICAS FINALES

VARIABLE	COMUNALIDAD	VARIANZA UNICA	FACTOR	% de VARIANZA EXPLICADA	% VARIANZA EXPLICADA ACUMULADA.
x_1	.87901	.12079	1	64.5	64.5
x_2	.84906	.15094	2	11.2	75.7
x_3	.53382	.46618			
x_4	.59765	.40235			
x_5	.77548	.22452			
x_6	.90710	.09290			

Tabla No. 9

En esta tabla No. 9 se observa que con sólo dos factores se explica el 75.7% de la variación total, teniendo el Factor 1 un porcentaje de la varianza explicada del 64.5%. Por lo que respecta a la rotación oblicua (oblimin) se obtuvieron los siguientes resultados.

ROTACION OBLICUA

VARIABLES	MATRIZ PATRON (PATTERN MATRIX) FACTOR 1	FACTOR 2	MATRIZ ESTRUCTURA (STRUCTURE MATRIX) FACTOR 1	FACTOR 2
x_1	.06613	**.89075**	.68008	**.93633**
x_2	.03566	**.89650**	.65357	**.92108**
x_3	-.05452	**.76714**	.47422	**.72956**
x_4	**.78448**	-.01668	**.77298**	.52402
x_5	**.85466**	.03705	**.88020**	.62613
x_6	**.95635**	-.00571	**.95241**	.65345

Tabla No. 10

Como podemos observar en esta tabla 10, la matriz patrón describe más claramente el agrupamiento de las variables manifiestas que la matriz estructura, la cual se refiere a los coeficientes de correlación entre cada variable y los factores 1 y 2. Sin embargo, un factor puede también contribuir a una variable dada indirectamente a través de otros factores correlacionados. Por lo tanto, la varianza total de una variable explicada *por un factor* no está dado por la suma de las contribuciones directas.

La comunalidad de una variable se refiere tanto a la contribución directa que es el caso de las cargas patrón que por ejemplo para x_1 puede ser expresada como .89075 del Factor 2 más .06613 del Factor 1, esto es, que .87901(.06613 x .68008 + .89075 x .93633) de la varianza de x_1 es explicada por la contribución conjunta de los 2 factores. Estas mezclas de conceptos son debidas a las correlaciones entre los factores, difícil de explicar desde el punto de vista estadístico.

En la tabla siguiente la No. 11 se presentan los resultados de las cargas factoriales para ambas rotaciones ortogonal y oblicua.

MATRICES DE CARGAS FACTORIALES (λ_{ik}), i=1,2,. . . ,6 k=1,2.

VARIABLES	MATRIZ ROTADA (VARIMAX)		MATRIZ PATRON (OBLICUA)	
	k_1(factor 1)	k_2(factor 2)	k_1(factor 1)	k_2(factor 2)
x_1	.39197	**.85169**	.06613	**.89075**
x_2	.36580	**.84572**	.03566	**.89650**
x_3	.23406	**.69212**	-.05452	**.76714**
x_4	**.72225**	.27568	**.78448**	-.01668
x_5	**.80737**	.35162	**.85466**	.03705
x_6	**.88592**	.34965	**.95635**	-.00571
	$\rho(k_1,k_2)$ = .93700		$\rho(k_1,k_2)$ = .68925	

Tabla No. 11

Anteriormente, cuando se habló del concepto de rotación en el capítulo 5 se comentó entre otras cosas, que los factores o variables latentes, en la realidad están correlacionados por lo cual la rotación oblicua proporciona frecuentemente factores significativos e interpretables. Precisamente en esta tabla se observa que los dos factores previamente determinados están más claramente definidos en la rotación oblicua que en la ortogonal, ya que en la rotación ortogonal hay cargas factoriales ≥ .30 para las variables manifiestas x_1, x_2, x_5 y x_6 en ambos factores. El primer factor es escencialmente una función de las variables manifiestas x_4, x_5, x_6 en tanto que el segundo factor es principalmente una función de las variables manifiestas x_1, x_2 y x_3 .

Estos mismos datos se procesaron con el paquete estadístico SAS versión 6.08 empleando el método de Máxima Verosimilitud con las rotaciones: a) ortogonal (varimax), b) ortogonal (equamax), c) ortogonal (orthomax) y d) ortogonal (quartimax) obteniendo los siguientes resultados:

La solución inicial sin rotar para cada uno de los factores, son los mismos valores obtenidos por el SPSS .

La varianza explicada de cada factor así como sus eigenvalores se presentan en la tabla siguiente. Estos valores son los mismos que se obtuvieron para las rotaciones Varimax, Equamax y Orthomax .

VARIANZA EXPLICADA Y EIGENVALORES DE DOS FACTORES

VARIABLE	COMUNALIDAD	FACTOR	VARIANZA
x_1	.878915	1	3.868300
x_2	.849150	2	.673836
x_3	.533846		
x_4	.597675		
x_5	.775547		
x_6	.907003		
TOTAL	4.542136		

TABLA No. 12

En la tabla siguiente se muestran las cargas factoriales para las rotaciones antes mencionadas (Varimax, Equamax y Orthomax) que fueron las mismas para las tres.

CARGAS FACTORIALES ROTACION ORTOGONAL

	FACTOR 1	FACTOR 2
x_1	.39202	**.85161**
x_2	.36578	**.84579**
x_3	.23403	**.69215**
x_4	**.72227**	.27567
x_5	**.80743**	.35157
x_6	**.88585**	.34968

TABLA No. 13

Comparando todos estos resultados con la rotación ortogonal obtenida por SPSS observamos que son prácticamente iguales . Sin embargo para el caso de la rotación QUARTIMAX las cargas factoriales y las varianzas explicadas de los factores rotados difieren ligeramente como puede apreciarse en la tabla siguiente:

CARGAS FACTORIALES ROTACION QUARTIMAX

	FACTOR 1	FACTOR 2
x_1	.44269	**.82640**
x_2	.41615	**.82217**
x_3	.27537	**.67677**
x_4	**.73759**	.23159
x_5	**.82718**	.30220
x_6	**.90533**	.29559

TABLA No. 14

Todas las rotaciones recuperan la estructura original de los datos, es decir, el Factor 1 describe las tres últimas variables y el Factor 2 las tres primeras variables, aunque también hay variables en el factor 1 con cargas ≥ .30, esto es, para x_1 y x_2 , y para el factor 2 las variables x_5 y x_6 .

Este ejemplo también fue procesado con el paquete LISREL ® 7 (1989) pero con el método de Análisis Factorial Confirmatorio, esto es, imponiendo restricciones tales como: 1) determinando qué pares de factores están correlacionados; 2) qué variables observadas están afectadas por qué o cuáles factores comunes; 3) qué variables observadas o manifiestas están afectadas, por un sólo factor y 4) qué pares de factores únicos están correlacionados.

Como se obtuvo anteriormente en el análisis factorial exploratorio una estructura de dos factores, é fue empleada para ser procesada por el LISREL 7 con el formato siguiente:

```
REPRODUCCION DEL ARTICULO DE  BARTHOLOMEW
DA NI=6 NO=100 MA=KM
LA
V1 V2 V3 V4 V5 V6
KM SY
1.0000
 .8636  1.0000
 .6792   .6738   1.0000
 .5099   .5164   .3274   1.0000
 .6347   .5738   .4165   .6751  1.0000
 .6396   .6232   .4614   .7377   .8384   1.0000
MO  NX=6  NK=2  PH=ST
LK
FACTOR1  FACTOR2
FR  LX(1,1)  LX(2,1)  LX(3,1)  LX(4,2)  LX(5,2)  LX(6,2)
OU ALL
```

Teniendo la siguiente estructura:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} \lambda_{11} & 0 \\ \lambda_{21} & 0 \\ \lambda_{31} & 0 \\ 0 & \lambda_{42} \\ 0 & \lambda_{52} \\ 0 & \lambda_{62} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \\ \delta_6 \end{pmatrix}$$

$$x = \Lambda y + \delta$$

Dando los siguientes resultados:

ESTIMACIONES INICIALES

Lambda x	Factor 1	Factor 2
v1	0.950	0.000
v2	0.915	0.000
v3	0.921	0.000
v4	0.000	0.777
v5	0.000	0.875
v6	0.000	0.952

Tabla No. 15

ESTIMACIONES (Máxima Verosimilitud)

Lambda x	Factor 1	Factor 2
v1	0.942	0.000
v2	0.918	0.000
v3	0.722	0.000
v4	0.000	0.772
v5	0.000	0.883
v6	0.000	0.950

Tabla No. 16

Así como la matriz Ψ (Phi) para la estimación máximo verosímil

MATRIZ PHI

	Factor 1	Factor 2
Factor 1	1.000	
Factor 2	0.716	1.000

Estos valores obtenidos nos proporcionan la siguiente matriz estructura (matrix structure) que es similar a la obtenida por el SPSS en la rotación OBLIMIN (oblicua).

	MATRIZ ESTRUCTURA SPSS F1	MATRIZ ESTRUCTURA SPSS F2	MATRIZ ESTRUCTURA LISREL F1	MATRIZ ESTRUCTURA LISREL F2
v1	.68008	**.93633**	**.9420**	.6745
v2	.65357	**.92108**	**.9180**	.6573
v3	.47422	**.72956**	**.7220**	.5170
v4	**.77298**	.52402	.5528	**.7720**
v5	**.88020**	.62613	.6322	**.8830**
v6	**.95241**	.65345	.6802	**.9500**

Tabla No. 17

Sólo que los factores están invertidos, esto es, el Factor 1 en el SPSS es el Factor 2 en LISREL y el Factor 2 en el SPSS es el factor 1 en LISREL.

Vale la pena señalar que Bartholomew concluye su artículo, diciendo que la rotación OBLICUA es bastante "*satisfactoria*" en recobrar la estructura del modelo originalmente propuesto.

PUNTAJES FACTORIALES (FACTOR SCORES)

Anteriormente se mencionó que el análisis factorial tiene, como principal objetivo expresar el conjunto de variables manifiestas a un grupo menor de variables llamadas latentes o simplemente factores. Este conjunto de factores o variables latentes puede emplearse fundamentalmente para dos propósitos:

1).-Para emplear las nuevas variables en análisis subsecuentes, por ejemplo, en el análisis de discriminante. Los puntajes incluídos son el resultado de todas las variables manifiestas que fueron consideradas importantes en la creación del nuevo factor, estos puntajes son conocidos con el nombre de *puntajes factoriales* (*factor scores*).

2).-Cuando el experimentador o analista está interesado en estimar los puntajes de y_i vectores del *i*-ésimo individuo de un factor dado. Por ejemplo, si en una investigación uno de los factores se ha llamado como *"habilidad verbal"*, resultará interesante estimar cada valor individual en este factor seleccionando una muestra de los sujetos que tuvieron los más altos puntajes en dicha prueba.

Los puntajes factoriales proporcionan la localización de cada observación en el espacio de los factores. A diferencia de componentes principales, en donde los puntajes de los componentes sí pueden calcularse directamente como combinaciones lineales. En análisis factorial, esto no es posible debido a que el modelo involucra la estimación de parámetros. Por ello los puntajes factoriales no pueden ser calculados directamente, pero en cambio pueden estimarse.

Los dos métodos de estimación de puntajes factoriales que más comúnmente se emplean y se expondrán en este trabajo, tienen una base intuitiva, proporcionando estimaciones que son funciones lineales de los datos observados. Supondremos que el vector observado de variables manifiestas para el i-ésimo individuo es x_i, y que la media de este vector muestral es $\bar{x}$. En función de este supuesto uno de estos métodos fue desarrollado por Bartlett citado por Krzanowski y Marriott (1995) empleando mínimos cuadrados en el modelo factorial lineal presentado en la ecuación (1) y buscando y_i para minimizar la siguiente expresión:

$$S = (x_i - \bar{x} - \hat{\Lambda}\, y_i)'\, \hat{\Psi}^{-1} (x_i - \bar{x} - \hat{\Lambda}\, y_i) \qquad (43)$$

Un cálculo sencillo produce la siguiente estimación:

$$\hat{y}_i = (\hat{\Lambda}'\hat{\Psi}^{-1}\hat{\Lambda})^{-1}\hat{\Lambda}'\hat{\Psi}(x_i - \bar{x}) \qquad (44)$$

El segundo método citado en Krzanowski y Marriott (1995) el que desarrolló Thomson (1951), en donde asumió normalidad de y y e en la expresión (1), colocando las variables
manifiestas y latentes en un sólo vector, empleando además las propiedades de la distribución normal multivariada para obtener el valor observado de y dado que $x=x_i$ como sigue:

$$E(y \mid x_i) = \Lambda'(\Lambda\Lambda' + \Psi)^{-1} (x_i - \mu) \qquad (45)$$

Un estimador obvio del vector de puntajes factoriales es el siguiente:

$$\hat{y}_i = \hat{\Lambda}'(\hat{\Lambda}\hat{\Lambda}' + \hat{\Psi})^{-1} (x_i - \bar{x}) \qquad (46)$$

Los puntajes factoriales generalmente son estandarizados (con media cero y varianza uno) antes de usarse con cualquiera de estos métodos de estimación, de acuerdo con los supuestos fundamentales del modelo. Ambos métodos proporcionan puntajes factoriales estimados como combinaciones lineales $\mathbf{A}(x_i - \bar{x})$ de media centrada para las variables manifiestas y la matriz **A** es frecuentemente llamada *matriz de estimaciones de puntajes factoriales* en las salidas de los paquetes estadísticos computacionales tales como el SPSS, donde

$$\mathbf{A} = \hat{\Lambda}'(\hat{\Lambda}\hat{\Lambda}' + \hat{\Psi})^{-1}$$

bajo el supuesto de normalidad de y y e.

Finalmente vale la pena mencionar que los puntajes o cargas factoriales son sólo aproximaciones a los factores y como tales son indicadores de errores a los que están propensos los factores subyacentes encontrados, por ello si el instrumento que se está empleando está validado y confiabilizado, muy probablemente el emplear los puntajes factoriales (*factor scores*) es la mejor alternativa, en caso que el instrumento carezca de confiabilidad y valídez o bien que sus valores obtenidos sean muy bajos, el emplear otras variables u otro modelo sería lo más recomendable.

CAPITULO 5
DOS EJEMPLO PRACTICOS.

A continuación se presenta en primer lugar la información referente a un ejemplo real sobre un instrumento de evaluación académica a nivel superior que consta de 21 reactivos en una escala de medición actitudinal tipo Lickert (1979). La escala tipo Lickert es una escala que corresponde a un nivel de medición ordinal, consistente en una serie de *reactivos* (*items* o *juicios*) ante los cuales se solicita la reacción del sujeto. El estímulo (*item* , *sentencia* o *reactivo*) que se presenta al sujeto representa la propiedad que el investigador está interesado en medir y las respuestas son solicitadas en términos de grados de acuerdo o desacuerdo, siempre o nunca que el sujeto tenga con el reactivo en particular.

El método del *summated ratings* de Lickert resulta de la suma algebráica de las respuestas de individuos a *reactivos* seleccionados previamente como válidos y confiables. Si bien la escala es aditiva, no se trata de encontrar reactivos que se distribuyan uniformemente sobre un contínuo "siempre - nunca" o "favorable - desfavorable", sino que el método de selección y construcción de la escala apunta a la investigación de reactivos que son definitivamente "siempre o nunca", "favorables o desfavorables", con relación al objetivo de estudio. El puntaje final del sujeto es interpretado como su posición en una escala de actitudes que expresa un contínuo con respecto al objeto de estudio.

Se aplicó a un total de 141 sujetos [1] en una escala que va de:

nunca casi nunca a veces casi siempre siempre

A continuación se presenta el instrumento, señalando entre paréntesis los reactivos *positivos* con el signo (+) y con el signo (-) los reactivos *negativos*.

(+) 1.- Su profesor imparte las 7 horas semanales asignadas a la materia.

(-) 2.- Se ha suspendido la clase por ausencia del profesor.

(+) 3.- El profesor inicia su clase a la hora señalada.

(+) 4.- Se cubre en clase el contenido temático señalado para cada examen departamental.

(+) 5.- Se resuelven en clase las dudas planteadas por los alumnos.

(+) 6.- El trato del profesor para el alumno es respetuoso.

(+) 7.- El profesor explica conceptos difíciles en forma clara.

(1) Los datos corresponden a un proyecto de investigación *estrictamente confidencial*, por lo cual no es posible dar más información.

(-) 8.- El profesor delega las clases en algún ayudante.

(-) 9.- El profesor acostumbra abandonar la clase mucho antes de la hora señalada.

(+) 10.- El profesor realiza evaluaciones periódicas a los temas vistos.

(+) 11.- El profesor notifica al alumno los resultados de dichas evaluaciones.

(+) 12.- El profesor favorece la participación de los alumnos en clase.

(+) 13.- El profesor motiva al alumno a resolver los problemas anatómicos.

(+) 14.- El profesor muestra interés en cada uno de sus alumnos.

(+) 15.- El profesor utiliza la Nómina Anatómica Internacional en sus clases.

(+) 16.- Durante el curso el profesor utilizó apoyos audiovisuales o didácticos.

(+) 17.- La bibliografía proporcionada por el profesor fue adecuada para el desarrollo del curso.

(+) 18.- Cuando surgían diferencias de opinión, el profesor permitía expresar las diversas opiniones.

(+) 19.- Durante las clases, se aprecia que el profesor domina la materia.

(+) 20.- Durante las clases, considera Ud. que apredende Anatomía.

(+) 21.- Por la Metodología de enseñanza que utiliza su profesor, considera que es un ejemplo a seguir.

Los reactivos positivos se calificaron de la forma siguiente:

nunca (1)
casi nunca (2)
a veces (3)
casi siempre (4)
siempre (5)

Por lo que respecta a los reactivos negativos, estos se calificaron como sigue:

nunca (5)
casi nunca (4)

a veces (3)
casi siempre (2)
siempre (1)

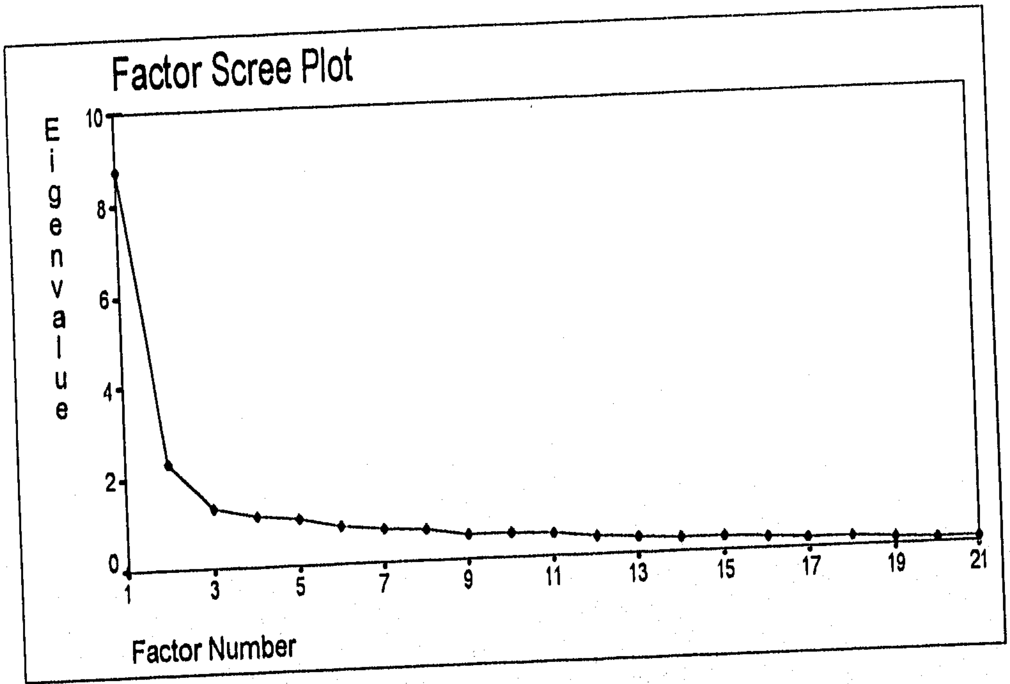
El procesamiento de esta información se llevó a cabo por medio del paquete estadístico SPSS . Se realizó primeramente un análisis de componentes principales, con el objeto de ayudar a determinar la extracción de los componentes que serían analizados en el Análisis Factorial, con los resultados siguientes.

ESTADISTICAS DEL ANALISIS DE COMPONENTES PRINCIPALES

FACTOR	EIGENVALOR	PORCENTAJE DE VARIANZA EXPLICADO	PORCENTAJE ACUMULADO DE VARIANZA EXPLICADO
1	8.78133	41.8	41.8
2	2.33436	11.1	52.9
3	1.33744	6.4	59.3
4	1.13448	5.4	64.7
5	1.02625	4.9	69.6
6	.84660	4.0	73.6
7	.76081	3.6	77.2
8	.68568	3.3	80.5
9	.56061	2.7	83.2
10	.53511	2.5	85.7
11	.48843	2.3	88.1
12	.39716	1.9	89.9
13	.37162	1.8	91.7
14	.32538	1.5	93.3
15	.31482	1.5	94.8
16	.26414	1.3	96.0
17	.23115	1.1	97.1

Tabla No.18

Además se obtuvo la gráfica de "*desmoronamiento*" (Scree Plot) donde puede observarse que a partir del factor 3 no se observan cambios "relevantes".



Gráfica No. 2

Teniendo en cuenta los tres criterios siguientes: 1) Porcentaje acumulado de varianza explicada de al menos un 60%; 2) Resultados de la gráfica de "*desmoronamiento*" ; 3) Principio de Parsimonia y estructura simple, se decidió trabajar con los tres primeros factores.

Posteriormente con estos tres factores se llevó a cabo un Análisis Factorial, empleando como método de extracción Componentes Principales (ver Kendall, 1980 pp. 52-53) con dos rotaciones: la ortogonal (VARIMAX) y la oblicua (OBLIMIN) eligiendo aquellos reactivos con cargas mayores o iguales a .30 con los resultados que se presentan en la tabla siguiente:

CARGAS FACTORIALES PARA LAS ROTACIONES ORTOGONAL Y OBLICUA

	ROTACION ORTOGONAL			**ROTACION OBLICUA**		
	F1	F2	F3	F1	F2	F3
R1	.15247	**.45260**	.44246	.04706	**.40093**	.41481
R2	.25589	.45098	**.54742**	.13868	.37481	**.50504**
R3	-.11316	**.74766**	.07204	-.04055	**.73423**	.09932
R4	.18952	**.61918**	.13974	.25507	-.03339	.19197
R5	**.81518**	-.03339	.19197	**.85059**	-.17015	.03079
R6	.36024	.41086	.37954	.28844	.33359	.30463
R7	**.69556**	.14027	.35141	**.67695**	.01058	.21757
R8	-.11661	.04744	**.75657**	.03179	.11372	**.71796**
R9	**.56068**	-.35884	.19237	**.55113**	-.24816	.27730
R10	.06230	**.79430**	-.07718	.02999	**.80986**	-.14319
R11	.26528	**.66724**	.20767	.20234	**.62491**	.12377
R12	**.68396**	.25373	.01206	**.72459**	.15860	-.15063
R13	**.78939**	.23207	.01895	**.83935**	.12047	-.16487
R14	**.74471**	.34558	.05048	**.77684**	.24021	-.12804
R15	.31985	.06663	**.71110**	.20127	-.04208	**.68939**
R16	**.62760**	.10660	.21721	**.63211**	-.00194	.09036
R17	**.72541**	.13498	.36535	**.70699**	-.00043	.22635
R18	**.81404**	.11736	.01026	**.87558**	.00038	-.17279
R19	**.59299**	.06466	.48808	**.54312**	-.06399	.39092
R20	**.79574**	.16166	.26510	**.80175**	.02556	.10196
R21	**.81391**	.22678	.27214	**.81575**	.08886	.10171

Tabla No. 20

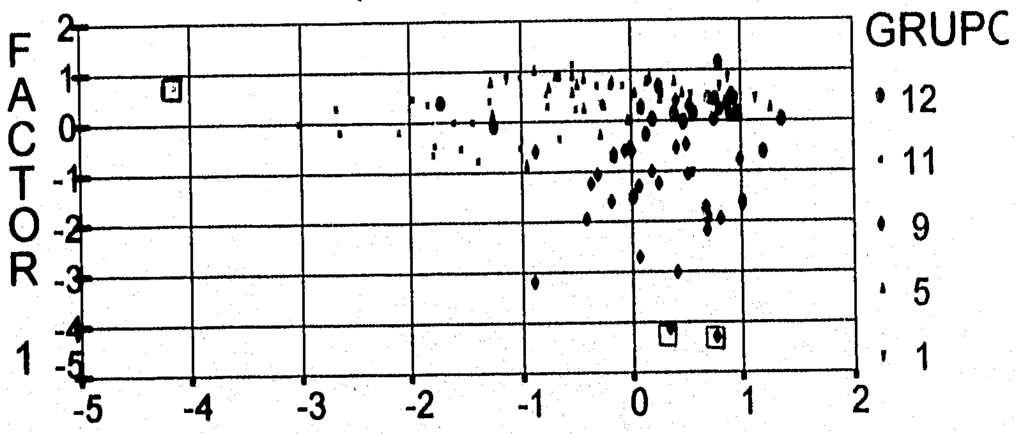
Los factores para ambas rotaciones en virtud de la similaridad, de las cargas factoriales y de los reactivos se llamaron de la siguiente manera: F1 "*Preparación Didáctica y Conocimiento de la materia*", F2 "*Cumplimiento Docente Institucional*" y F3 "*Responsabilidad Docente* ". Particularmente el reactivo 6 (R6) como resultado de la rotación ortogonal sería conveniente eliminarlo, ya que sus cargas factoriales en los Factores 1 2 y 3 son bastante similares, lo cual nos indica que dicho reactivo está evaluando aspectos diferentes a lo que originalmente se planeó, es decir, no discrimina en tal sentido. Sin embargo en la rotación oblicua sucede para este reactivo algo similar, decidiendo eliminarlo.

De los resultados anteriormente obtenidos se sugiere ampliar la temática de los reactivos en los factores encontrados así como la inclusión de otros temas, ya que como hemos podido observar en el Factor 1 que explica aproximadamente el 42%, las cargas factoriales significativas (mayores o iguales a .30) corresponden a 12 reactivos de 20 (puesto que uno fue eliminado), esto es, el 60% se incluyen aspectos didácticos y de conocimiento de la materia. Para el Factor 2 con un porcentaje de explicación del 11.1% se incluyen 5 reactivos, es decir, el 25% y para el Factor 3 con un porcentaje de explicación bajo, esto es, del 15% se incluyen los 3 reactivos restantes.

Aspectos de evaluación académica, tales como: estrategias de exposición; conducción de la clase; relación clase-ejercicio profesional; organización y preparación de la clase; empleo de auxiliares didácticas; relación con los estudiantes y procedimientos para evaluar el aprendizaje entre otros, sería conveniente considerarlos en la reestructuración de dicho instrumento.

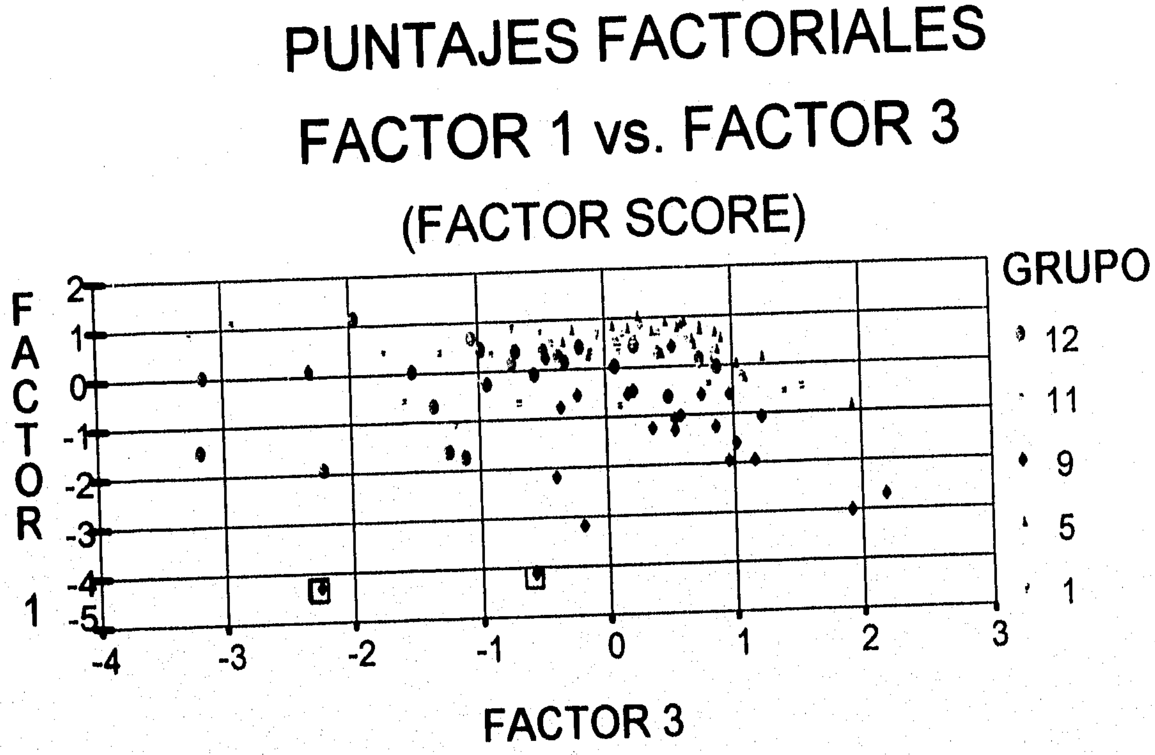
También se obtuvieron las siguientes gráficas de los puntajes factoriales obtenidos del método de rotación VARIMAX de los tres factores extraídos en función del grupo de pertenencia : F1 vs. F2 gráfica No. 3 , F1 vs. F3 gráfica No. 4 y F2 vs. F3 gráfica No. 5. Observándose en la primera de ellas dos grupos de observaciones claramente diferenciados, integrado por los grupos No. 11 y No. 5 fundamentalmente, y para el segundo los grupos No. 9 y No. 12. Por lo que se refiere al F1 con el F3, se observa la formación de un sólo grupo de observaciones para el grupo No. 9 y para la gráfica del F2 con el F3, se observa la formación de un grupo de observaciones integrado por el grupo No. 11 .

PUNTAJES FACTORIALES
FACTOR 1 vs. FACTOR 2
(FACTOR SCORE)



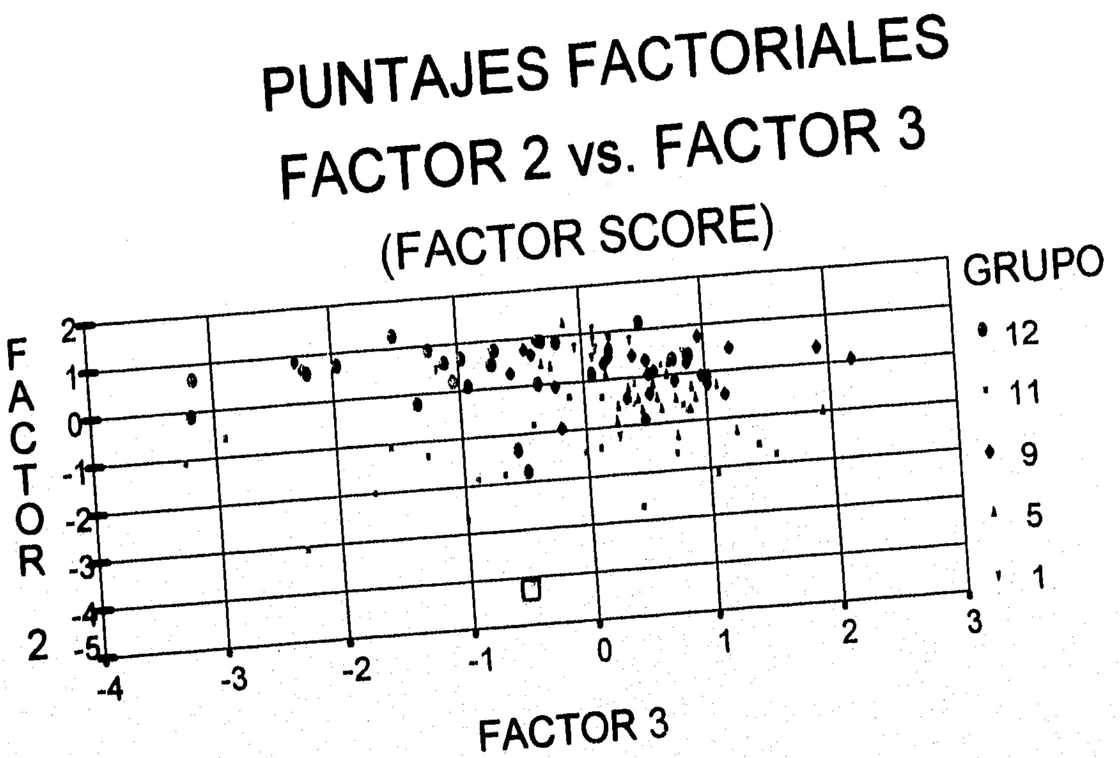
GRAFICA No. 3

□ observación potencialmente *aberrante* (outlier).



GRAFICA No. 4

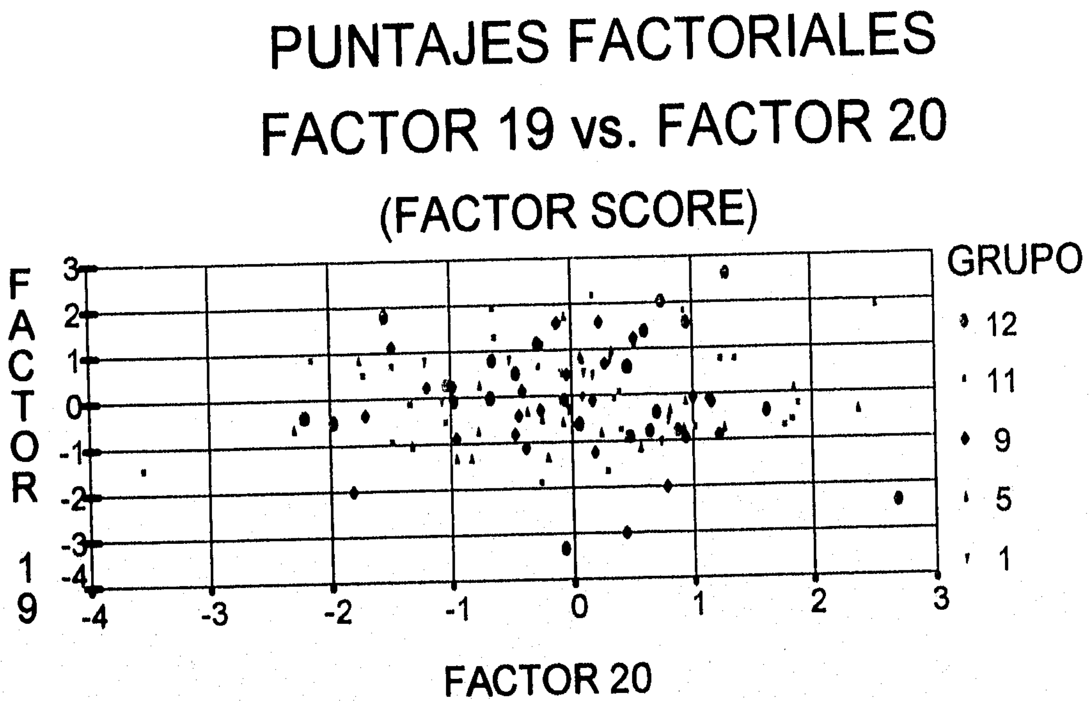
☐ observación potencialmente *aberrante* (outlier).



GRAFICA No.5

□ observación potencialmente *aberrante* (outlier).

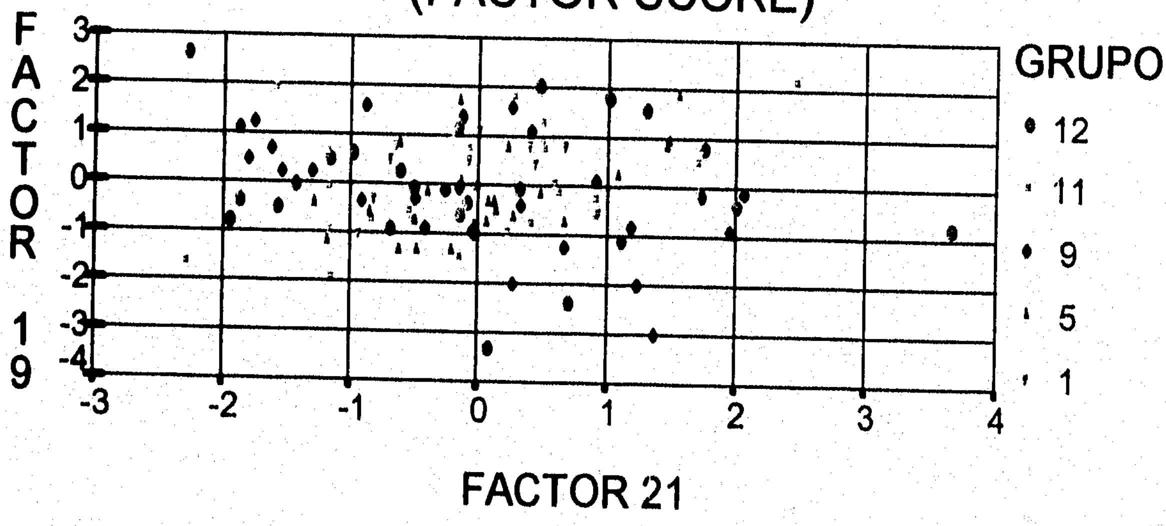
Por otro lado, con el objeto de estudiar la posible presencia de "puntos aberrantes" (outliers) se graficaron las cargas de las 3 últimas componentes del Análisis de Componentes Principales: C19 vs. C20 gráfica No. 6 , C19 vs. C21 gráfica No. 7 y C21 vs. C22 gráfica No.8 y tal y como podemos apreciar no hay puntos aberrantes que nos hayan llevado a la necesidad de analizarlos con detalle y tal vez decidir su eliminación.



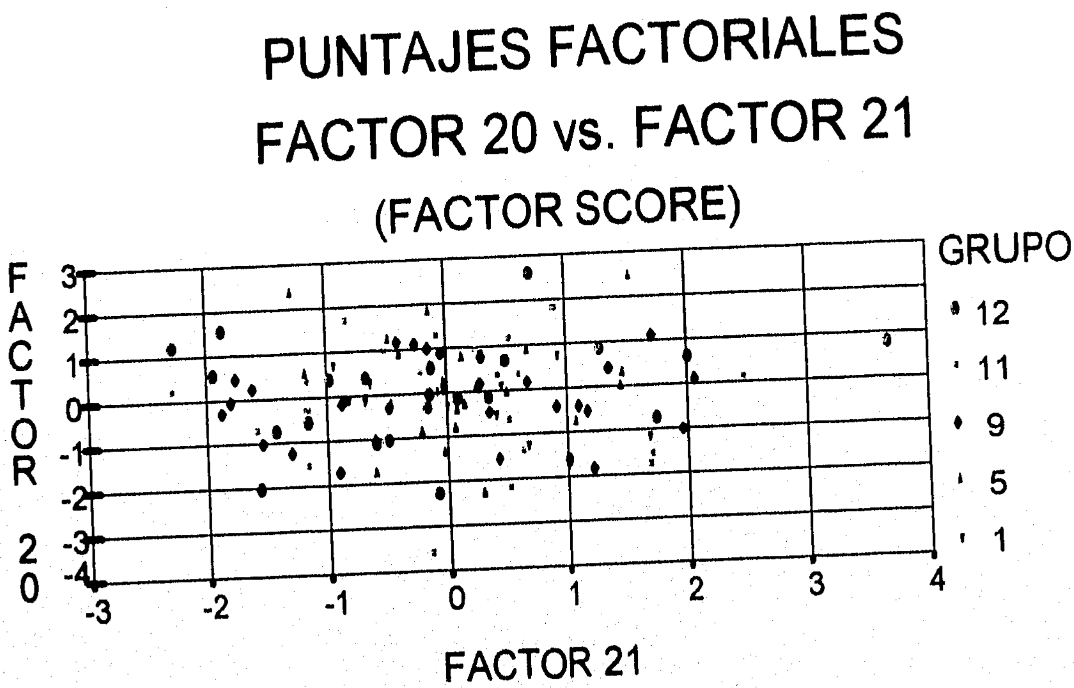
GRAFICA No. 6

Sin embargo en las gráficas No. 3 , No. 4 y No. 5 se aprecia en cada una de ellas alguna(s) observación considerada como potencialmente *aberrante*. Al eliminarse y nuevamente hacer los análisis correspondientes, los resultados fueron similares a los obtenidos sin dicha eliminación.

PUNTAJES FACTORIALES
FACTOR 19 vs. FACTOR 21
(FACTOR SCORE)



GRAFICA No. 7


PUNTAJES FACTORIALES
FACTOR 20 vs. FACTOR 21
(FACTOR SCORE)
GRUPO
12
11
9
5
1
FACTOR 20
FACTOR 21
3
2
1
0
-1
-2
-3
-4
-3
-2
-1
0
1
2
3
4

GRAFICA No. 8

Por lo que respecta al segundo ejemplo, este se incluyó debido a que el anterior ejemplo de evaluación enseñanza-aprendizaje de la Facultad de Medicina solamente se contó con una muestra de n=141 sujetos y 21 reactivos. Esta muestra resultó insuficiente para procesarla con el programa PRELIS, el cual envió un mensaje indicando que el tamaño de muestra era pequeño para el cálculo de la estimación de la matriz de covarianza asintótica (no basada en la distribución normal). Particularmente en el manual del paquete PRELIS (1986, pp. 2.28 y 3.32) se señala que para poder realizar el cálculo antes mencionado, así como las covarianzas y correlaciones de datos a nivel ordinal (esto es, correlaciones policóricas) se requiere al menos un tamaño de muestra de n=200 para $p < 12$ y para valores de $p \geq 12$ será de $1.5p(p+1)$, donde p es el número de variables observadas o manifiestas. A continuación se reproduce la tabla dada por el manual del PRELIS.

Número de variables	Tamaño de muestra mínimo.
3	200
5	200
10	200
15	360
20	630
25	975
30	1395

Con el fin de ilustrar el uso del paquete PRELIS y LISREL 7 presentamos este segundo ejemplo el cual fue realizado por la Coordinación de Matemáticas y Estadística de la Facultad de Psicología UNAM, Flores L. y Cuevas C. (1995). En este trabajo se investigaron las estrategias de estudio en Matemáticas y Estadística empleadas por los estudiantes de los cuatro primeros semestres de la licenciatura, en los que se cursan dichas materias. De un total de 21 grupos, se seleccionaron aleatoriamente 10 de ellos, esto es, el 48% del total de la población. En cada grupo se aplicó el cuestionario a todos los alumnos de cada grupo, obteniéndose una muestra de n=343. El instrumento consta de catorce reactivos cuyo nivel de medición es ordinal con tres opciones de respuesta como se muestra a continuación: *lo realizas*, *algunas veces* y *no lo realizas*.

REACTIVOS
1.- Subrayado
2.- Resúmenes
3.- Fichas de trabajo
4.- Cuadros sinópticos
5.- Formularse preguntas
6.- Parafraseo
7.- Ejercicios
8.- Corrección de ejercicios y exámenes

9.- Examen preliminar
10.- Acordeones
11.- Formularios
12.- Lectura activa
13.- Tutorial
14.- Analogías (dos o más cosas semejantes).

Como una primera aproximación al desarrollo de un modelo, estos 14 reactivos se procesaron en el paquete SPSS empleando componentes principales y rotación oblicua para cuatro factores, seleccionados en base a la taxonomía de estrategias de aprendizaje de Weinstein y Mayer (1986). Se escogió este método de rotación debido a que se supuso que los factores están correlacionados. Se obtuvieron los siguientes resultados.

CARGAS FACTORIALES DE LA MATRIZ ESTRUCTURA.

FACTOR STRUCTURE

	F1	F2	F3	F4
V_1	.05368	**.40819**	.21629	-.07621
V_2	.23388	.39120	**.50635**	-.05357
V_3	.18363	.20362	**.43509**	.18952
V_4	.25658	.08793	**.51632**	.12652
V_5	.18510	.01918	.10448	**.43517**
V_6	.05496	.03570	.06589	**.30206**
V_7	**.94674**	-.03314	-.03398	.16594
V_8	**.42982**	.08752	.28491	.23171
V_9	.24432	.02341	.28265	**.30418**
V_{10}	.00367	.03957	.26663	.00989
V_{11}	**.33538**	.23018	.28926	.06945
V_{12}	.05284	**.70293**	-.05505	.43323
V_{13}	-.05484	.12108	-.12751	.25063
V_{14}	.13470	.06360	.07765	**.45159**

Tabla No. 21

COMUNALIDADES Y EIGENVALORES.

VAR.	COMUNALIDAD	FACTOR	EIGENVALOR
V1	.20636	1	1.27134
V2	.37849	2	1.30473
V3	.22879	3	.88752
V4	.28504	4	.57600
V5	.20473		
V6	.09360		
V7	.99900		
V8	.23625		
V9	.17966		
V10	.07680		
V11	.18472		
V12	.65641		
V13	.10225		
V14	.20758		

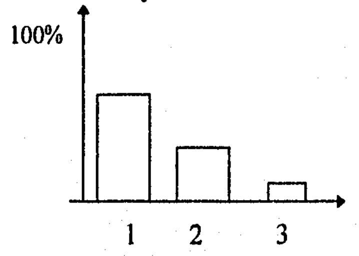
Tabla No. 22

Para estos cuatro factores se obtuvo alrededor de sólo un 29% de explicación de la varianza total.

Con estos resultados iniciales se procedió a trabajar con PRELIS y LISREL 7. Para este ejemplo el valor de n=343 resultó suficiente ya que el valor requerido por PRELIS fue de 1.5(14)(14 + 1) = 315. Desde luego existen otros valores sugeridos para n por otros autores y paquetes, por ejemplo el paquete EQS (pronunciado "X") citado por Potthast M. (1993) sugiere una razón de n:p de al menos 10:1 para distribuciones libres y probablemente mayor si se desea que los errores estándar y los valores de Ji-cuadrada sean correctos (es decir, no estén sub o sobre valuados). Finalmente Muthen (1989) señala que al menos se deben tener 1000 observaciones cuando se use el Método para Variables Categóricas (CVM) en el Análisis Factorial Confirmatorio (CFA) para diez variables dicotómicas y quien en el año de 1987 dio a conocer una metodología para abordar este tipo de problemas con datos categóricos, llamada *Metodología para Variables Categóricas (CVM)* conjuntamente con un paquete llamado LISCOMP (Muthen y Kaplan 1992).

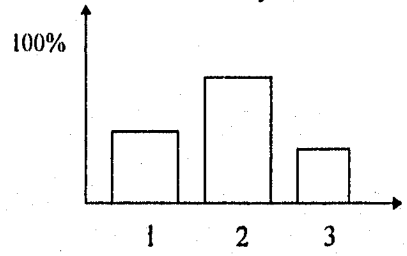
Señalado por los autores Muthen y Kaplan (1992) y Potthast M. (1993) es conveniente llevar a cabo un análisis descriptivo de las variables bajo estudio, con el objeto de "detectar" variables ordinales con sesgos y curtosis relativamente "altas" que nos pueden llevar a resultados incorrectos en la estimación de modelos tales como probabilidad de sesgos negativos en los errores estándar y valores de Ji-cuadrada sobrestimados . Las implicaciones de esto son por un lado que los parámetros pueden aparecer como significativos cuando realmente no lo son y por el otro que los modelos pueden considerarse pobremente ajustados a los datos reales, cuando en realidad son correctos. La versión del PRELIS 2 empleada en este ejemplo lleva a cabo este análisis presentándose los resultados a continuación.

E1: Subrayado



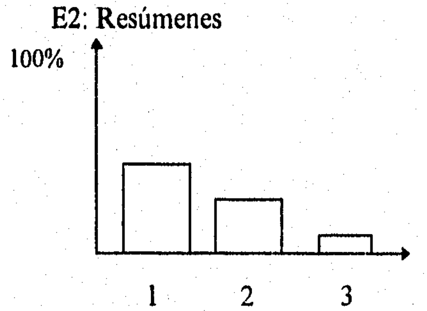
Media = 1.397
Varianza = .380
Sesgo = 1.303
Curtosis = .595

E3: Fichas de trabajo



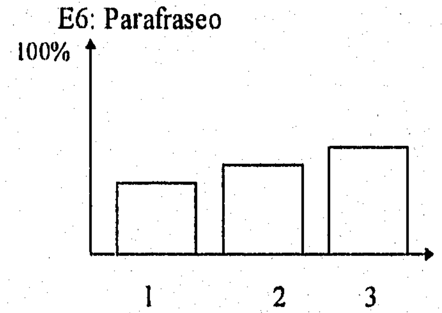
Media= 1.927
Varianza = .471
Sesgo = .094
Curtosis = -.871

E2: Resúmenes



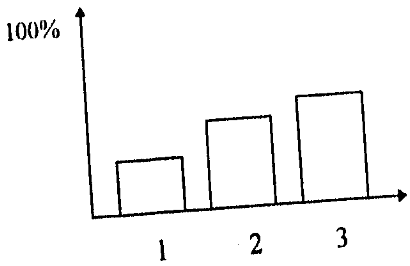
Media = 1.519
Varianza = .350
Sesgo = .651
Curtosis = -.529

E6: Parafraseo



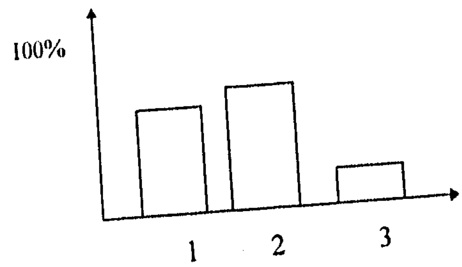
Media = 2.166
Varianza = .630
Sesgo = -.306
Curtosis = -1.351

E3: Fichas de trabajo



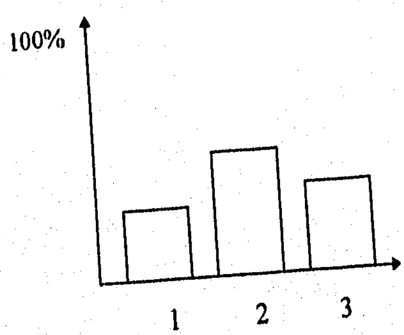
Media = 2.329
Varianza = .502
Sesgo = -.570
Curtosis = -.855

E7: Ejercicios



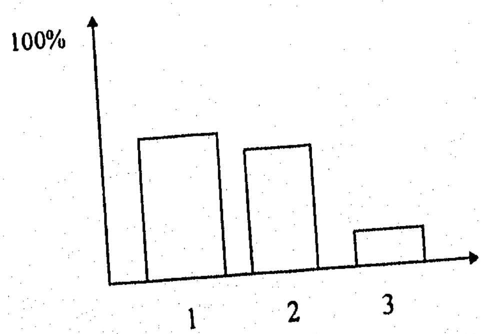
Media = 1.638
Varianza = .389
Sesgo = .443
Curtosis = -.658

E4: Cuadros Sinópticos



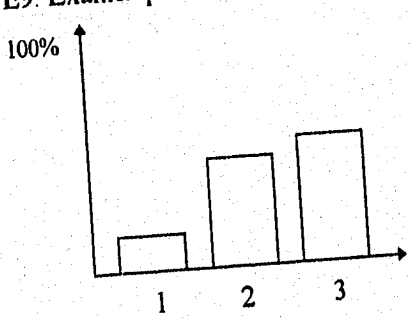
Media = 2.137
Varianza = .528
Sesgo = -.215
Curtosis = -1.084

E8: Corrección de ejercicios y exámenes

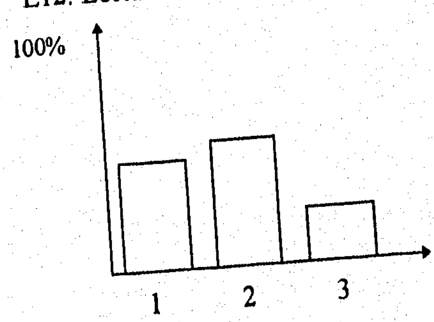


Media = 1.624
Varianza = .452
Sesgo = .616
Curtosis = -.684

E9: Examen preliminar

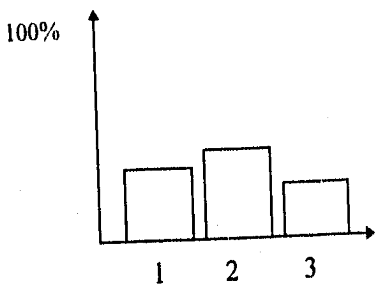


E12: Lectura activa



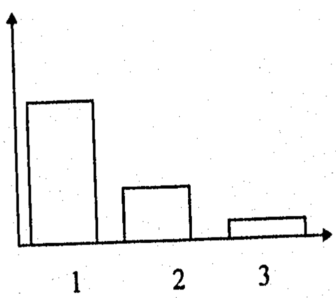
Media = 2.338
Varianza = .482
Sesgo = -.567
Curtosis = -.801

E10: Acordeones



Media = 1.974
Varianza = .634
Sesgo = .047
Curtosis = -1.419

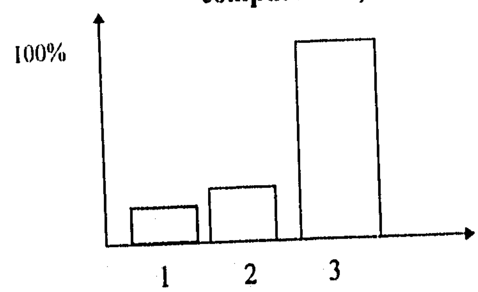
E11: Formularios



Media = 1.411
Varianza = .424
Sesgo = 1.324
Curtosis = .510

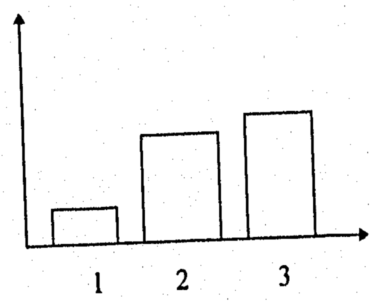
Media = 1.743
Varianza = .513
Sesgo = .423
Curtosis = -.974

13: Tutorial (sistemas expertos en computadora).



Media = 2.828
Varianza = .213
Sesgo = -2.742
Curtosis = 6.856

E14: Analogías



Media = 2.332
Varianza = .474
Sesgo = -.543
Curtosis = -.796

De los resultados anteriores se observa que once variables tienen asimetría < 1, y siete de ellas con signo negativo. En cuanto a la curtosis díez son < 1 y ocho de ellas negativas. Particularmente llama la atención la variable 13 (estrategia: *Tutorial*) ya que presenta un sesgo y curtosis muy altas de -2.742 y 6.856 respectivamente. De manera empírica era algo ya esperado debido a la muy escasa producción de material didáctico por computadora en estas áreas así como el limitado acceso a ellos y/o al uso de computadoras. Otra variable que se observó podría presentar algunos problemas en este modelo fue la estrategia 10: *acordeones*, ya que su distribución es muy cercana a la uniforme (no discrimina) y es la que tiene el sesgo mayor pero con signo negativo. Sin embargo, inicialmente fueron incluidas en el PRELIS y posteriormente en el LISREL 7, dándonos cargas factoriales estimadas muy bajas, es decir, de .200 y de .192 para las estrategias 14 y 10 respectivamente. Por ello fueron eliminadas del modelo final que se presenta a continuación:

ESTRUCTURA MATRICIAL DEL MODELO: ESTRATEGIAS DE ESTUDIO EN MATEM.

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9 \\ x_{11} \\ x_{12} \\ x_{14} \end{pmatrix} = \begin{pmatrix} 0 & \lambda_{12} & 0 & 0 \\ 0 & 0 & \lambda_{23} & 0 \\ 0 & 0 & \lambda_{33} & 0 \\ 0 & 0 & \lambda_{34} & 0 \\ 0 & 0 & 0 & \lambda_{54} \\ 0 & 0 & 0 & \lambda_{64} \\ \lambda_{71} & 0 & 0 & 0 \\ \lambda_{81} & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_{94} \\ \lambda_{11,1} & 0 & 0 & 0 \\ 0 & \lambda_{12,2} & 0 & 0 \\ 0 & 0 & 0 & \lambda_{14,4} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \\ \varepsilon_7 \\ \varepsilon_8 \\ \varepsilon_9 \\ \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon_{14} \end{pmatrix}$$

Las instrucciones dadas al PRELIS para este modelo se presentan a continuación:

```
ESTRATEGIAS DE ESTUDIO EN MATEMÁTICAS FAC. DE PSICOLOGÍA
DA NI=12
LA
E1 E2 E3 E4 E5  E6 E7 E8 E9 E10 E11 E12
RA FI=A:LIS.DAT
OR ALL
OU MA=PM SM=A:LISI2.PML SA=A:LISI2.ACP PA
```

La línea OR ALL señala que todas las variables se encuentran a nivel ordinal. La línea OU señala que la matriz de correlaciones policóricas es guardada en el archivo LISI2.PML y la matriz de covarianza asintótica es guardada en el archivo LISI2.ACP

La matriz de correlaciones *policóricas* obtenida por PRELIS se presenta a continuación:

MATRIZ DE CORRELACIONES POLICORICAS

	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12
E1	1.000											
E2	.380	1.000										
E3	.136	.372	1.000									
E4	.086	.306	.414	1.000								
E5	.026	.008	.135	.177	1.000							
E6	.010	.066	.068	.013	.264	1.000						
E7	-.038	.085	.054	.124	.204	.042	1.000					
E8	.079	.266	.165	.243	.203	-.059	.452	1.000				
E9	.040	.182	.167	.194	.199	.177	.216	.342	1.000			
E10	.244	.266	.173	.310	-.025	.094	.341	.222	.116	1.000		
E11	.283	.195	.168	.014	.192	.151	.044	.114	.093	.184	1.000	
E12	-.025	-.026	.152	.046	.239	.243	.143	.240	.157	.184	.247	1.000

Tabla No. 23

PRELIS también calculó la matriz de covarianza asintótica y con ambas matrices se empleó el LISREL 7 para el modelo anteriormente presentado de cuatro factores dentro del Análisis Factorial Confirmatorio (CFA) en donde las variables (estrategias) corrección de exámenes y ejercicios (E8), ejercicios (E7) y formularios (E11) son asumidos por el Factor 1 denominado *Estrategias Optimas*; las variables subrayado (E1) y lectura activa (E12) son asumidas en el Factor 2 llamado *Estrategias de Comprensión*; las variables resúmenes (E2), fichas de trabajo (E3) y cuadros sinópticos (E4) son asumidas por el Factor 3 denominado *Estrategias de Organización y Síntesis*; finalmente las variables: formularse preguntas (E5), parafraseo (E6), examen preliminar (E9) y analogías (E14) son asumidas por el Factor 4 llamado *Estrategias poco empleadas*.

Posteriormente se procesó este modelo con las dos matrices previamente calculadas en el paquete LISREL con las siguientes instrucciones:

```
ESTRATEGIAS DE ESTUDIO EN MATEMÁTICAS FAC. DE PSICOLOGÍA
DA NI=12 NO=343 MA=PM
LA
E1 E2 E3 4 E5 E6 E7 E8 E9 E10 E11 E12
PM FI=A:LISI2.PML
AC FI=A:LISI2.ACP
MO NX=12 NK=4 PH=ST
FR LX 7 1 LX 8 1 LX 10 1 LX 1 2 LX 11 2 LX 2 3 LX 3 3 LX 4 3
FR LX 5 4 LX 6 4 LX 9 4 LX 12 4
0U SE
```

La línea AC implica que el método de estimación WLS (mínimos cuadrados ponderados) será usado por omisión en lugar del método ML (máxima verosimilitud). Los resultados se presentan a continuación .

ESTADÍSTICAS FINALES . Valores de $\hat{\lambda}_{pk}$ (mínimos cuadrados ponderados)

	F1	F2	F3	F4
V1	0.000	0.476	0.000	0.000
V2	0.000	0.000	0.612	0.000
V3	0.000	0.000	0.613	0.000
V4	0.000	0.000	0.584	0.000
V5	0.000	0.000	0.000	0.471
V6	0.000	0.000	0.000	0.320
V7	0.551	0.000	0.000	0.000
V8	0.702	0.000	0.000	0.000
V9	0.000	0.000	0.000	0.537
V11	0.513	0.000	0.000	0.000
V12	0.000	0.580	0.000	0.000
V14	0.000	0.000	0.000	0.462

Tabla No. 24

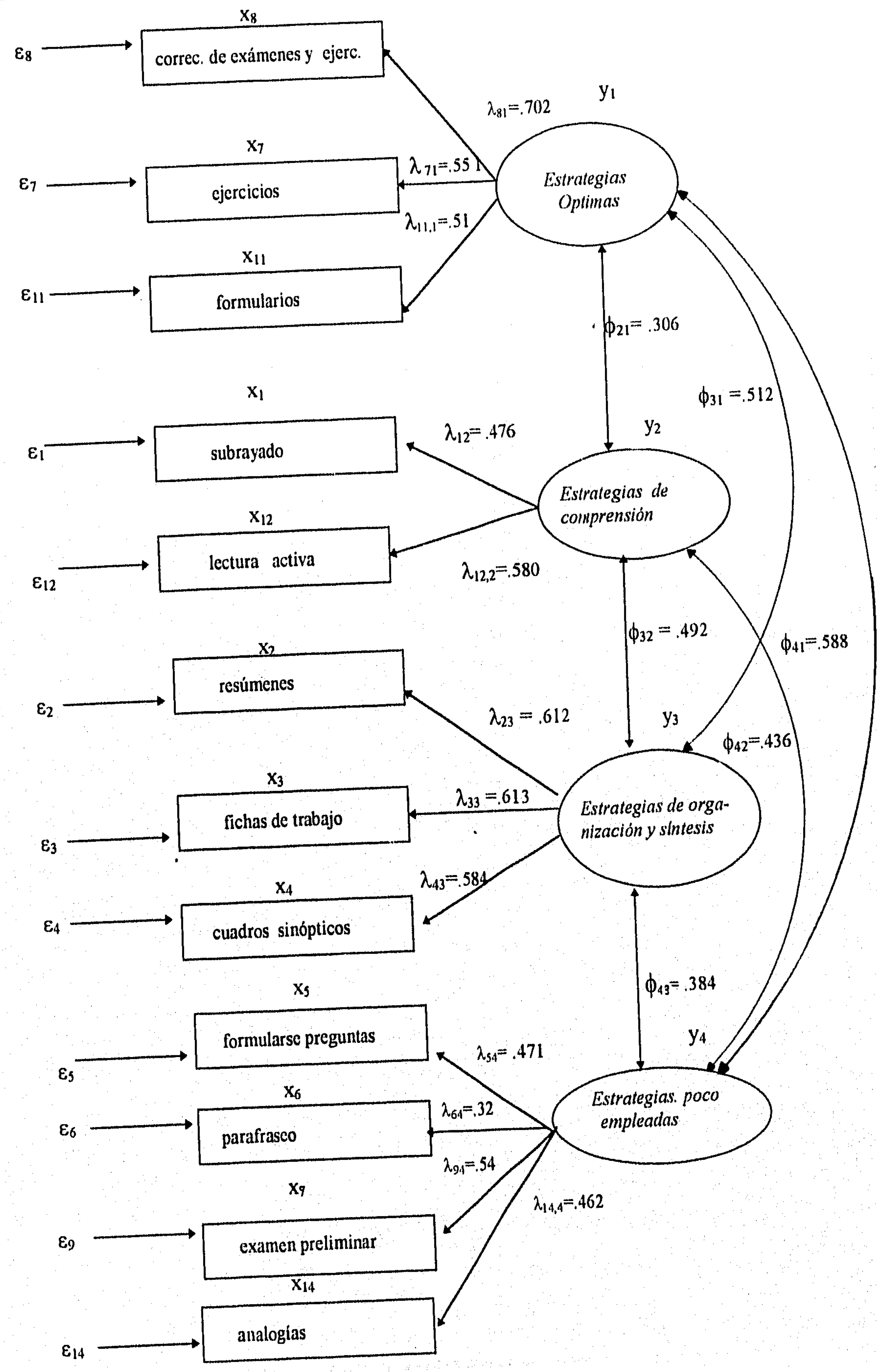
VALORES $\hat{\phi}_k$ (correlación entre los factores).

	F1	F2	F3	F4
F1	1.000			
F2	0.306	1.000		
F3	0.512	0.492	1.000	
F4	0.588	0.436	0.384	1.000

Tabla No. 25

Con esta información se elaboró el siguiente diagrama.

ESTRUCTURA DEL MODELO DE ESTRATEGIAS DE ESTUDIO EN MATEMATICAS



Particularmente Jöreskog y Sörbom (1988, pp. 195-196) señalan que emplear el método de estimación WLS (mínimos cuadrados ponderados) con la matriz de correlación producto momento de Pearson es incorrecto debido a que estas correlaciones están basadas en puntajes normales y las correlaciones de variables que no cumplen con este supuesto (como podrían ser las ordinales) tienen sesgos.

Algunos otros resultados se comentan a continuación. LISREL 7 obtuvo las correlaciones múltiples al cuadrado (SMC) para cada una de las doce variables manifiestas consideradas en este modelo x_i las cuales se refieren a la cantidad de varianza en cada una de ellas, que es explicada por las cuatro variables latentes (factores) y_i conjuntamente:

CORRELACIONES MÚLTIPLES AL CUADRADO (SMC)

E1 = .227
E2 = .374
E3 = .376
E4 = .341
E5 = .221
E6 = .103
E7 = .304
E8 = .493
E9 = .288
E11= .263
E12= .336
E14= .213

Tabla No. 26

Siendo la variable (estrategias) *corrección de exámenes y ejercicios* la de mayor porcentaje de varianza, seguida por la variable (estrategias) *fichas de trabajo* y por la variable (estrategia) *resúmenes*.

Con el fin de investigar la *validez concurrente* de las estrategias de estudio en matemáticas se usó como criterio el número de veces que se ha cursado la materia. Para ello sería conveniente realizar un análisis de discriminante de dos grupos con los puntajes factoriales obtenidos. Uno de estos grupos estaría formado por los alumnos que cursan por primera vez la materia y el otro por los alumnos que la han cursado más de una vez. Como una aproximación a este análisis se graficaron los puntajes factoriales obtenidos en el SPSS con los dos grupos, no se hizo con LISREL 7 ya que no proporciona los puntajes factoriales para cada sujeto, solamente en forma matricial. Esto debido a que trabaja con la matriz de correlación policórica y la matriz de varianza covarianza asintótica.

Se obtuvieron seis gráficas de los cuatro factores. En ellas no se observa la formación de dos grupos que difieran en las cuatro estrategias de estudio latentes. Estos resultados pueden explicarse por el hecho de que dentro del grupo de los alumnos que están cursando la materia y tendrán que recursarla, y por lo tanto poseen características similares a las del otro grupo. Consecuentemente los grupos de criterio adecuados serían un grupo de aprobados y otro de *reprobados* en la materia.

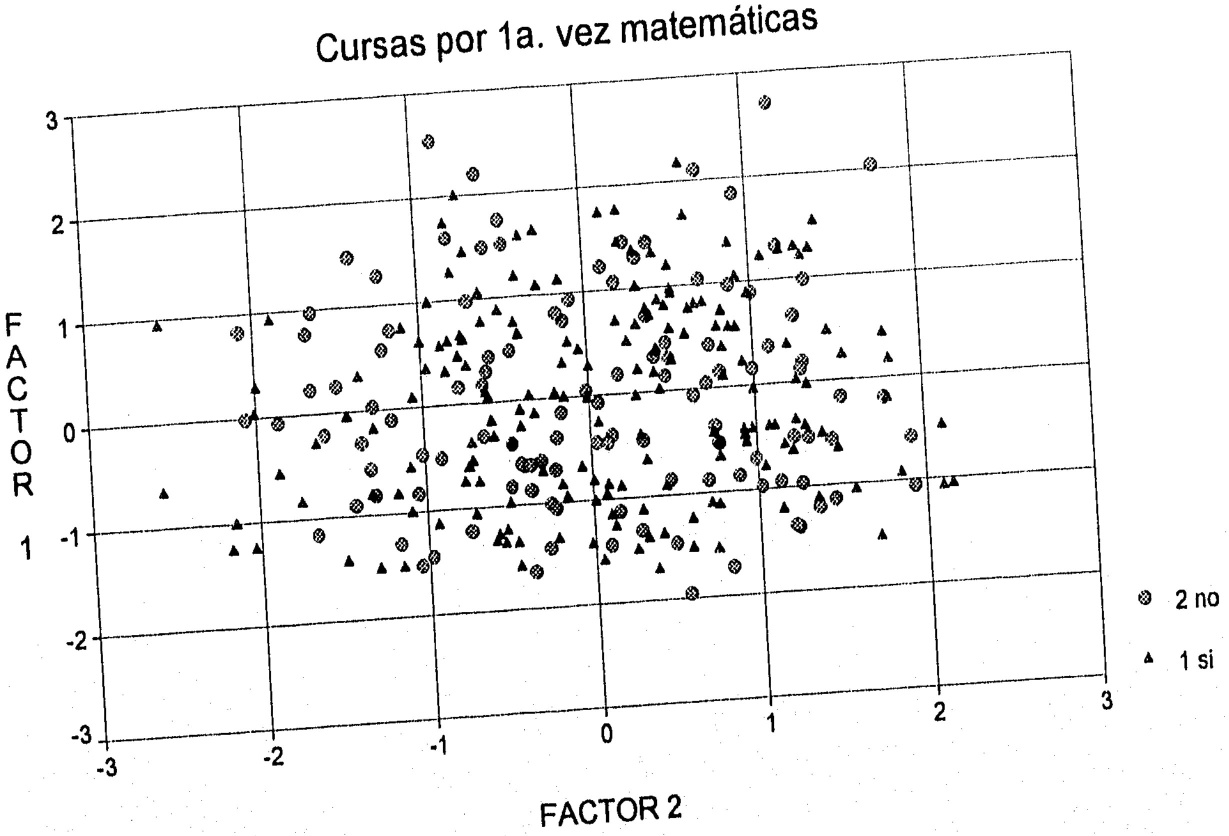
Otros resultados interesantes que proporciona LISREL 7 se presentan a continuación. El coeficiente total de determinación para las doce variables observadas fue de 0.950 el cual nos señala que las cuatro dimensiones del modelo son buenas. El valor que se obtuvo de la Ji-cuadrada fue de 94.01 con gl=49 y un valor de P <.00001 lo cual nos indica de manera formal que Σ (matriz modelo) y **S** (matriz de covarianza muestral) difieren en forma significativa. A este respecto Jöreskog y Sörbom (1988 p. 43) enfatizan que el uso de la Ji-cuadrada como prueba del modelo no es válida en muchas aplicaciones, ya que en la gran mayoría de trabajos empíricos, el modelo es sólo tentativo y por lo tanto considerado como una aproximación a la realidad. Por lo tanto el uso de esta estadística quedará determinado por los datos y situación con la cual trabaje el investigador. En este ejemplo se tuvo como punto de partida un modelo empírico (debido a las pocas investigaciones que se han hecho en esta área y a este nivel en la propia Facultad e inclusive en otros lugares) el cual constituye una primera aproximación. Este modelo es susceptible de mejorarse sugiriendo una reformulación el cuestionario empleado.

Por lo que respecta a los índices GFI (*Goodness of Fit Index*), AGFI (*Adjusted Goodness of Fit Index*) diremos que los valores obtenidos fueron de .963 y .941 respectivamente, los cuales son mayores a .9 (estos índices toman valores entre cero y uno, esto significa un modelo pobre para el primer valor; cero y un modelo perfecto para el segundo valor; uno) lo cual nos indica que tenemos un modelo empírico bueno .

A continuación se presentan las seis gráficas de los cuatro factores y los puntajes factoriales, observándose además de lo anterior que no hay presencia de puntos potencialmente aberrantes (*outliers*).

PUNTAJES FACTORIALES

FACTOR 1 vs. FACTOR 2

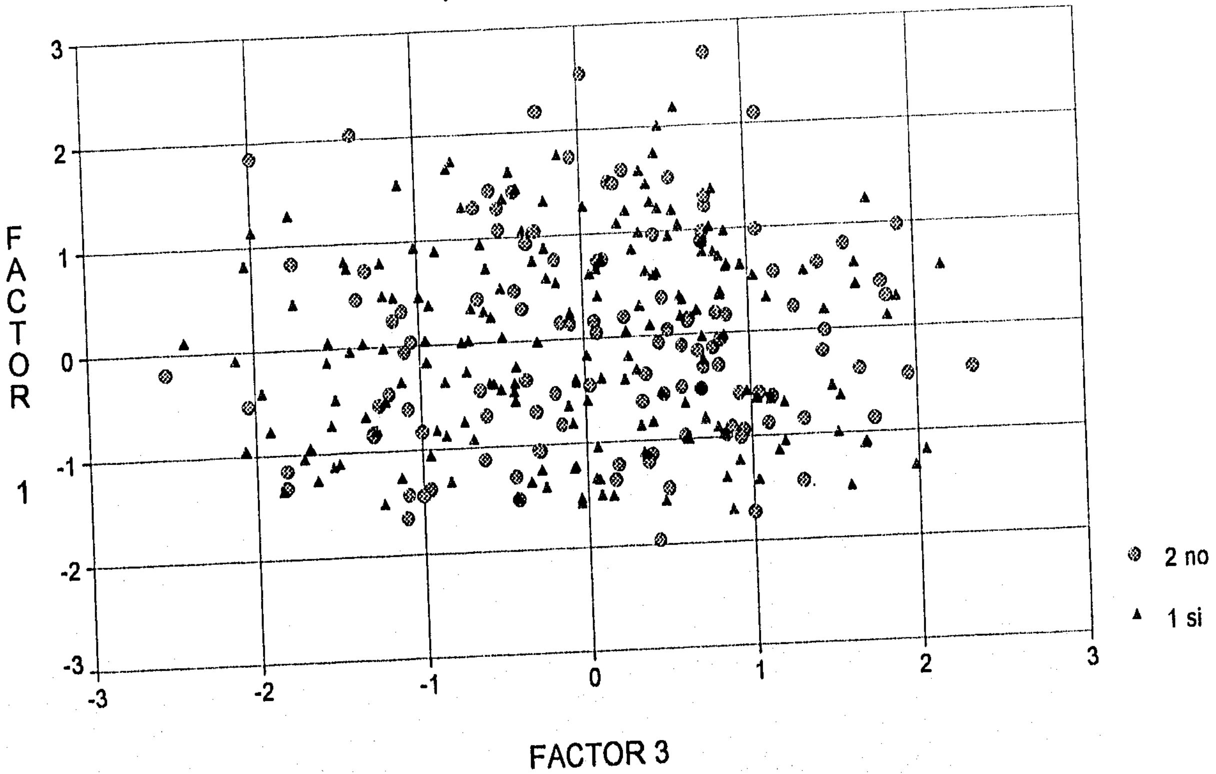


GRAFICA No. 9

PUNTAJES FACTORIALES

FACTOR 1 vs. FACTOR 3

Cursas por 1a. vez matemáticas

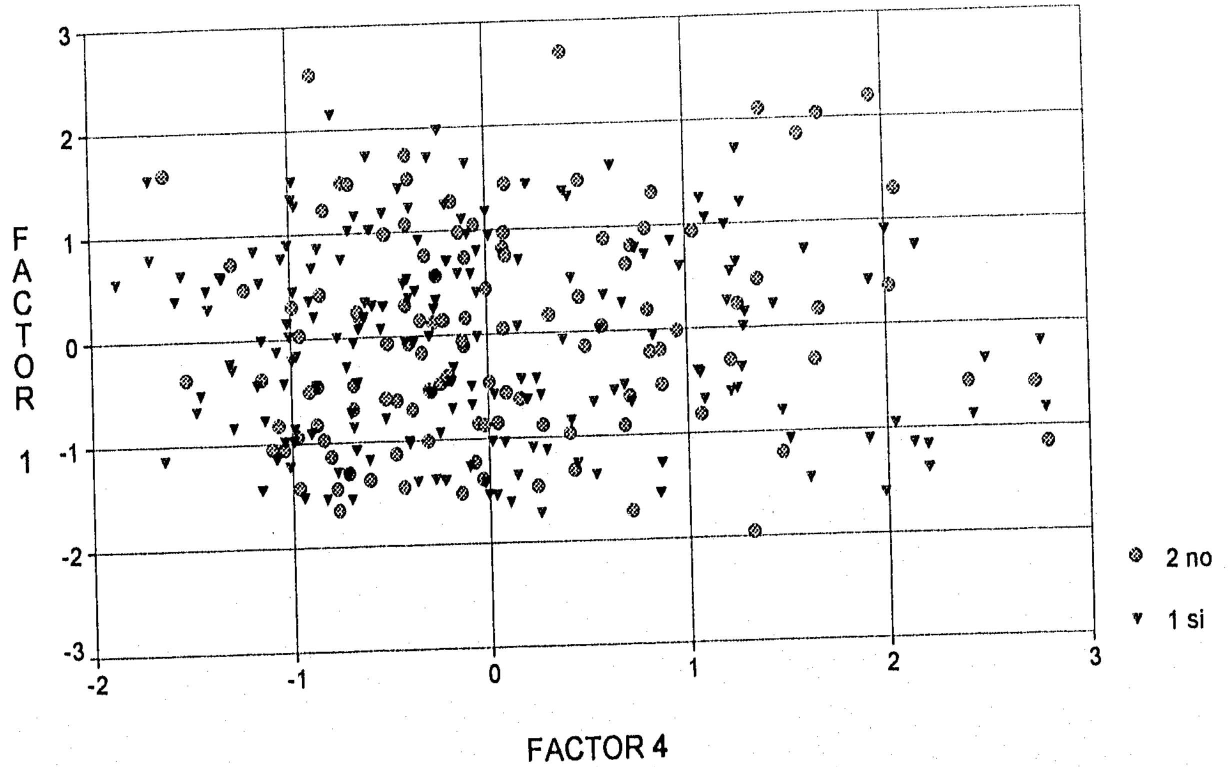


GRAFICA No. 10

PUNTAJES FACTORIALES

FACTOR 1 vs. FACTOR 4

Cursas por 1a. vez matemáticas

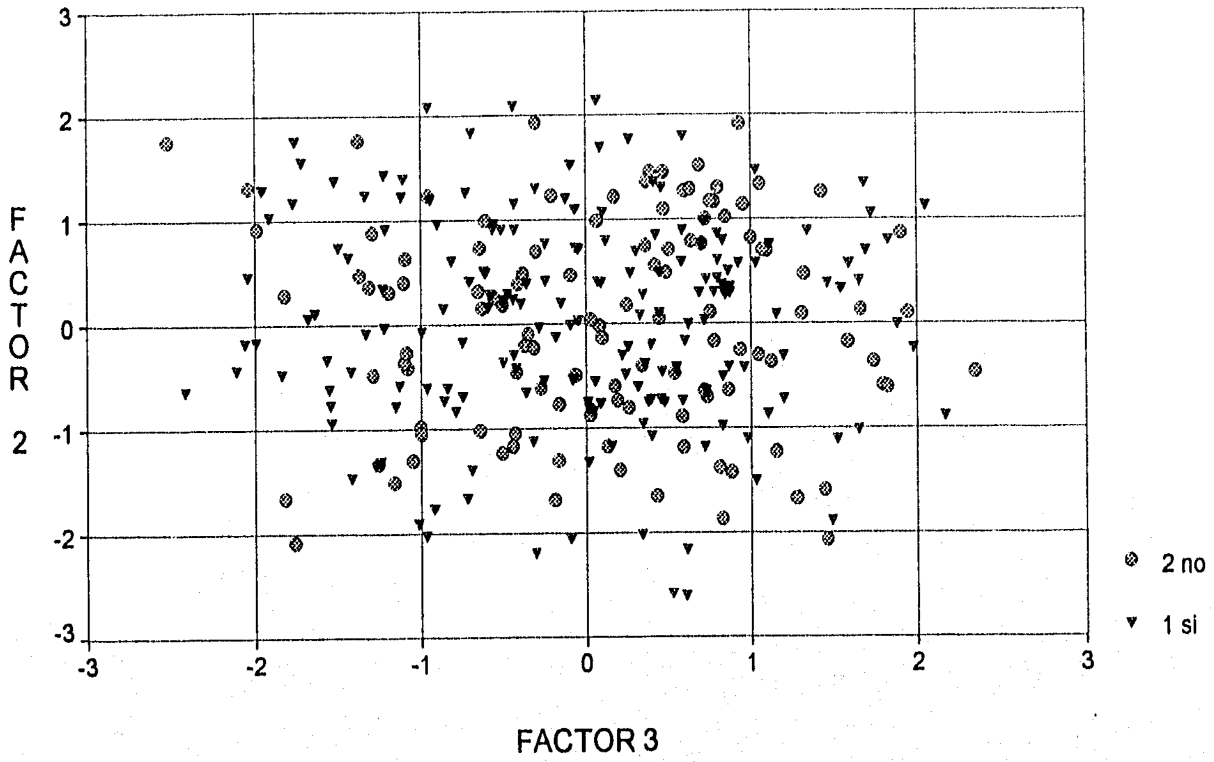


GRAFICA No. 11

PUNTAJES FACTORIALES

FACTOR 2 vs. FACTOR 3

Cursas por 1a. vez matemáticas

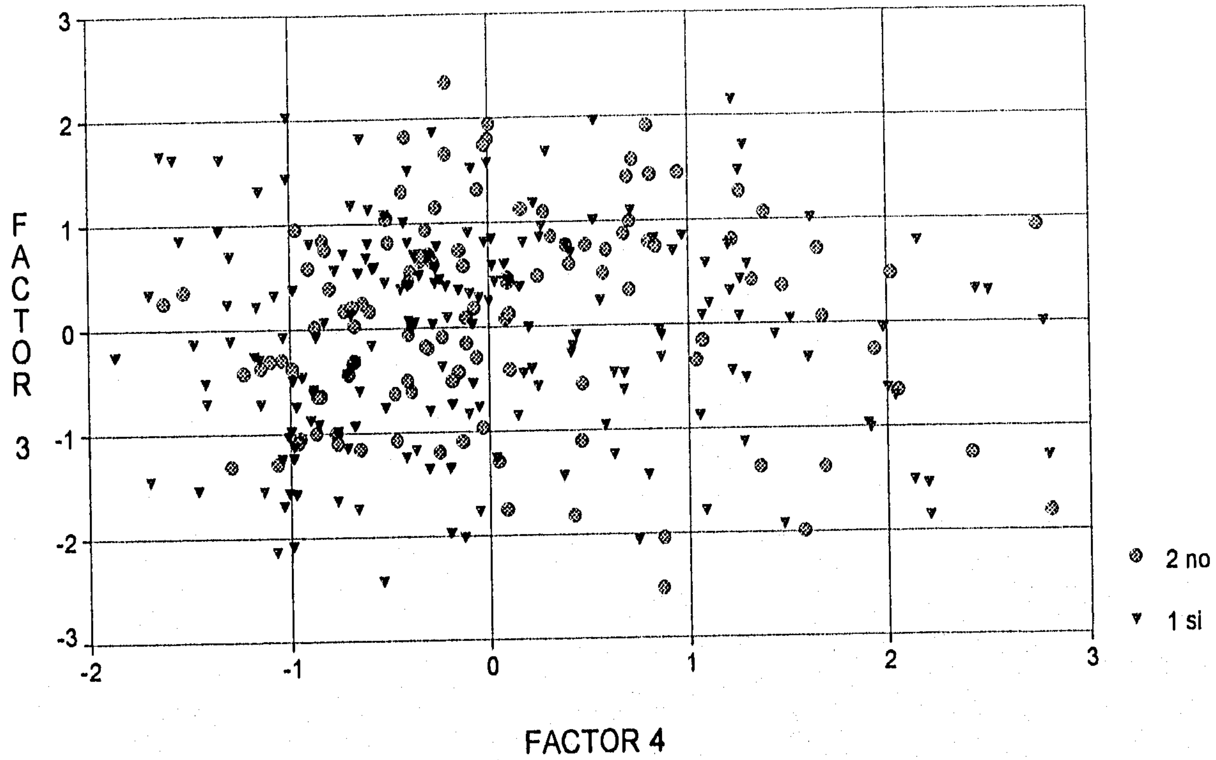


GRAFICA No. 12

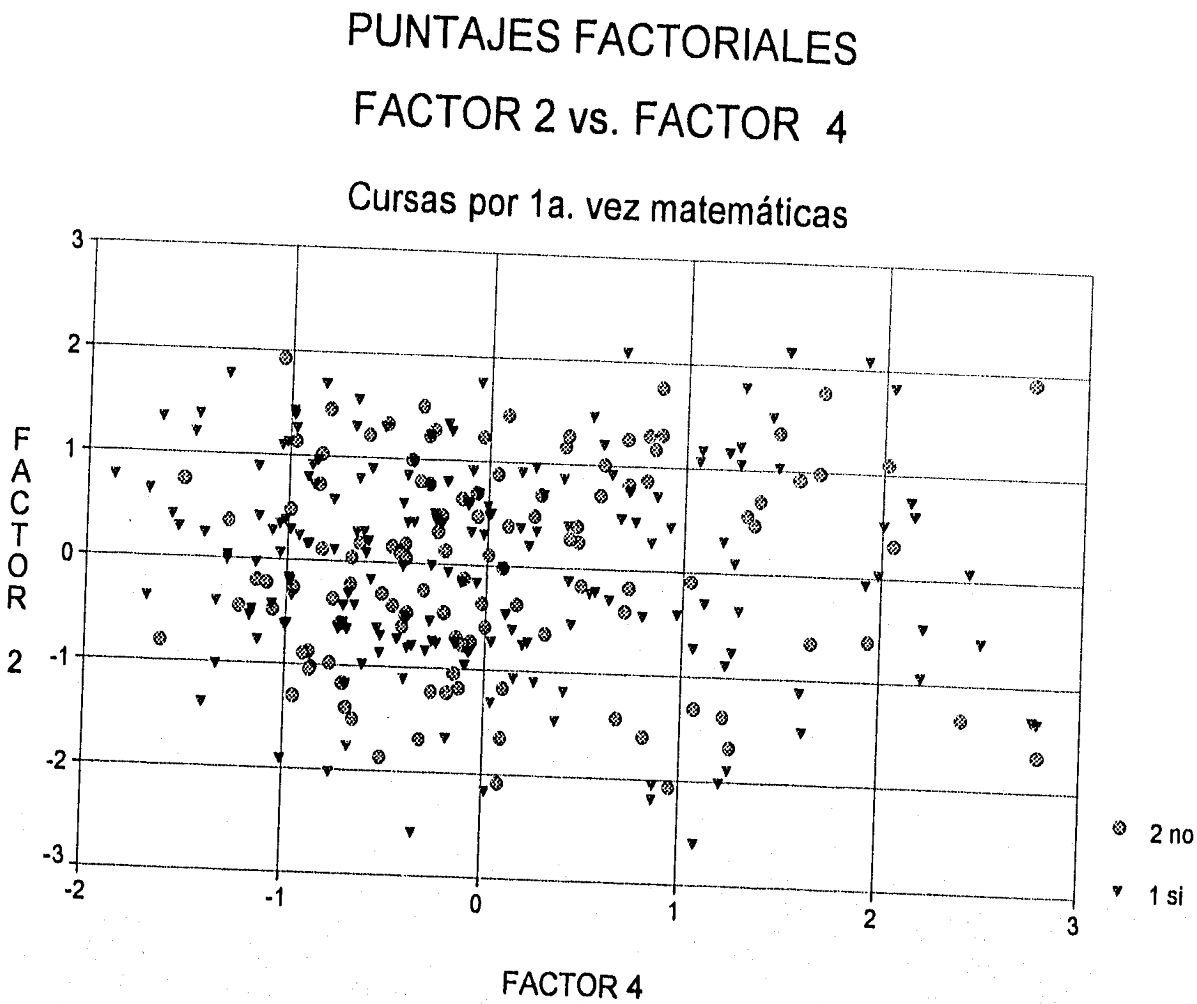
PUNTAJES FACTORIALES

FACTOR 3 vs. FACTOR 4

Cursas por 1a. vez matemáticas



GRAFICA No. 13



GRAFICA No. 14

CAPITULO 6

CONCLUSIONES

El Análisis Factorial ha ido paulatinamente ocupando un lugar importante en el campo de las Ciencias Sociales, particularmente en Psicología. Una revisión de las publicaciones más importantes en los campos nacionales e internacionales de los últimos cinco años, reportan un total de 256 artículos enfocándose la gran mayoría de ellos al área de la Psicometría. De esta lista, sólo uno de ellos fue realizado por investigadores nacionales y fue publicación nacional (Díaz Loving y Colb ,1989). Aunque cabe mencionar que el Dr. M. Choynowski de la Universidad Pedagógica Nacional, ha trabajado y publicado en investigación sobre la *Agresión en adolescentes mexicanos* usando análisis factorial (M. Choynowsky, 1988). Esta información refleja en parte la necesidad de que los estadísticos trabajen de manera conjunta con los psicólogos particulamente en el área de Psicometría, pues resulta un hecho inegable que un buen número de psicólogos no sólo en el área laboral sino también en la investigación hacen frecuentemente uso de esta técnica multivariada, pero ¿cómo lo hacen?, creo que los resultados anteriores reflejan en buena parte la respuesta a esta interogante: de manera incorrecta y desconocida.

Es una realidad que el Análisis Factorial es considerado por muchos estadísticos un particular tipo de modelo con bases fuertes en el área específica de su aplicación. De ahí el poco interés por la gran mayoría de los estadísticos con excepción quizás de estadísticos renombrados como Barttlet, M. Kendall , D. Bartholomew, K. G. Jöreskog y B. Muthen. Sin embargo cada vez un mayor número de profesionales de las áreas sociales y económicas incursionen en esta área con resultados poco satisfactorios debido en buena parte a la falta de un trabajo interdisciplinario. Hay quienes como señala Hope (1982) consideran que el Análisis Factorial es similar al método científico medieval ya que es la naturaleza quien dicta las explicaciones y por lo tanto es considerado por muchos un método *heurístico*. Sin embargo para la Psicología como para otras Ciencias, un gran número de trabajos donde la *validación concurrente, predictiva o de contenido* es generalmente muy difícil de realizarse con resultados exitosos, el Análisis Factorial surge como una herramienta poderosa para la *validación de constructo (o de concepto)* al igual que el Análisis de Correspondencia.

Este último tipo de validación se emplea, cuando no existe un criterio externo, claro y evidente en la interpretación. Entendiendo por un concepto o constructo el atributo o carcateristica que se postula para los individuos y que habrá de reflejarse a través de los resultados obtenidos en un test o instrumento de medición (Cronbach y Meehl, 1955).

No obstante lo anterior, el Análisis Factorial tiene su "talón de Aquiles", el cual radica en el hecho de que un mismo problema puede tener *n* estructuras diferentes y todas ellas "*válidas*", además de que no hay "*rigidez*" ni en el diseño ni en las hipótesis salvo para procedimientos muy específicos y particularmente cuando se emplea como técnica estadística exploratoria. Salvo los estadísticos renombrados anteriormente como M. Barttlet y M. Kendall en los inicios y últimamente de K.G. Jöreskog , D. Bartholomew y B. Muthen quienes le han dado mayor rigor al análisis factorial.

Aunque por otro lado D. Bartholomew y K.G. Jöreskog enmarcan al análisis factorial en modelos más generales e igualmente con bases sólidas: el primero en los modelos de variables latentes (*Latent Variables Models*) y el segundo en modelos estructurales (*Structural Equations Models*).

Dentro de los modelos estructurales latentes a continuación se presenta una tabla de clasificación en función del nivel de medición de las variables (Bartholomew, 1987, p 4.):

METODOS DE CLASIFICACION PARA VARIABLES LATENTES.

		VARIABLES MANIFIESTAS	
		CUANTITATIVAS	CATEGORICAS
VARIABLES LATENTES	CUANTITATIVAS	ANALISIS FACTORIAL (*FACTOR ANALYSIS*)	ANALISIS CARACTERISTICO LATENTE. ANALISIS FACTORIAL DE DATOS CATEGORICOS (*LATENT TRAIT ANALYSIS FACTOR ANALYSIS OF CATEGORICAL DATA*).
	CATEGORICAS	ANALISIS DE PERFIL LATENTE (*LATENT PROFILE ANALYSIS*)	ANALISIS DE CLASE LATENTE (*LATENT CLASS ANALYSIS*)
		ANALISIS MIXTOS (*ANALYSIS OF MIXTURES*)	

TABLA No. 24

La defensa que puede hacerse al Análisis Factorial es en el sentido de que la experimentación rigurosa frecuentemente es de interés académico, ya que sólo cuando se está en un laboratorio donde "casi" todo está controlado se alcanzan estos niveles de rígidez y "cientificidad" , careciendo de *validez externa*. La realidad no cabe en un laboratorio es muchísimo más compleja especialmente la conducta humana y el Análisis Factorial ha mostrado ser y lo sigue haciendo una técnica *heurística* útil para aclarar y ordenar las ideas de quienes hacen uso de ella. Particularmente para las Ciencias Sociales, está ocurriendo un cambio de paradigma (Méndez R. Y. 1989) que se refleja en el uso cada vez mayor de la Estadística y las Matemáticas, reconociendo dos aspectos básicos: 1) no es posible ser totalmente objetivo y 2) la naturaleza no es representativa del conocimiento obtenido.

Antes de concluir este capítulo considero importante señalar algunos aspectos sobre los paquetes LISREL 7 y SAS.

Particularmente para el paquete SAS resultó un tanto desconcertante el que la rotación OBLICUA no estuviera incluída en la Subrutina *Factor Analysis*. En cambio ofrece al usuario un buen número de rotaciones ortogonales tal y como anteriormente se señaló. Quien merece un buen número de comentarios fue el paquete LISREL 7, debido en gran parte a que sus resultados no son estrictamente comparables con los paquetes SPSS y SAS por las causas que a continuación se presentan:

a) LISREL calcula el Análisis Factorial exclusivamente del tipo confirmatorio al imponer directamente restricciones en los coeficientes de los factores -entre otras cosas- , es decir, decidiendo "*a prori*" qué variables manifiestas deben figurar en qué factores, asignando a las restantes variables un coeficiente de cero, obteniéndo de este modo estructuras simples o parsimoniosas, resultanto un método con el espíritu Kantiano, es decir, de "someter a la naturaleza a la inquisición", imponiéndole condiciones artificiales para obtener la respuesta a hipótesis preconcebidas. Mientras que los paquetes SAS y SPSS obtienen estructuras simples al hacer rotaciones.

b) Particularmente para los ejemplos que en este trabajo se desarrollaron no se logró obtener una solución de carácter exploratorio en el lenguaje de Jöreskog, ya que no se pudo obtener una solución con cargas factoriales "*libres*", es decir, sin restricciones.

Probablemente esto sucedió debido a las dos situaciones siguientes:

1) La documentación contenida en los manuales, aunque hace referencia al modelo con el que se trabajó, le dedica cuando mucho dos cuartillas con información sumamente escueta y aunque las instrucciones son seguidas paso a paso, el paquete no obtiene dicho modelo y 2) lo anterior pienso no es posible, es decir, esta situación es improbable por lo cual es casi seguro que no posea los conocimientos necesarios para poderlo trabajar.

La versión empleada en este trabajo fue como ya se señaló LISREL 7, sin embargo se tuvo acceso a la última versión al final de la elaboración de este trabajo: LISREL 8 y PRELIS2 (SSI Scientific Software Interenational, Chicago, IL 60615 USA). Legándose a los mismos resultados que con la versión LISREL 7 y PRELIS 1. Los avances de la última versión son en el sentido de proporcionar gráficos (*Path Analysis*) y el estadístico de prueba de la Normal multivariada, entre otros.

LISREL es un paquete mucho más general que lo que el Análisis Factorial implica. No en el sentido de diversidad de técnicas estadísticas como las que proporcionan los paquetes SAS y SPSS, sino en cuanto al alcance de ecuaciones estructurales, en donde el Análisis Factorial es sólo un *submodelo*, llamado por Jöreskog *submodelo 1*(Jöreskog 19898, p. 75).

Con la experiencia práctica que se obtuvo, parece que LISREL 7 es uno de los más potentes paquetes del Análisis Factorial Confirmatorio de un total de siete que existen actualmente en el mercado (Waller, Niels G. 1995), pero justamente para el modelo confirmatorio, por lo cual es necesario como usuario del mismo tener conocimientos más profundos de las herramientas matemáticas que subyacen al modelo de ecuaciones estructurales en general.

ANEXO A
SPEARMAN
INSTRUCCIONES DEL SPSS/PC+

```
set disk='a:spe1.res'.
data list free matrix /x1 to x6.
begin data.
1.00
.83 1.00
.78 .67 1.00
.70 .67 .64 1.00
.66 .65 .54 .45 1.00
.63 .58 .51 .51 .40 1.00
end data.
factor read=correlation triangle/variables=x1 to x6/
analysis=x1 to x6/print=all/
extraction=pc/
rotation=norotate/
rotation=varimax.
factor read=correlation triangle/variables=x1 to x6/
analysis=x1 to x6/format=sort/
criteria=factors(1)/plot=eigen/extraction=pc/
rotation=norotate/print=all.
```

EJEMPLO DE BARTHOLOMEW.
SPSS /PC +

```
set disk='a:bart.res'.
data list free matrix/x1 to x6.
begin data.
1.0000
.8636 1.0000
.6792 .6738 1.0000
.5099 .5164 .3274 1.0000
.6347 .5738 .4165 .6751 1.0000
.6396 .6232 .4614 .7377 .8384 1.000
end data.
factor read=correlation triangle/variables=x1 to x6/
analysis=x1 to x6/print=all/
criteria=factors(2)/plot=eigen/
extraction=ml/rotation=varimax/plot=rotation(1,2).
factor read=correlation triangle/variables=x1 to x6/
analysis=x1 to x6/print=all/
criteria=factors(2)/plot=eigen/
extraction=ml/rotation=oblimin/plot=rotation(1,2).
```

SAS
INSTRUCCIONES

```
title analisis factorial. Bartholomew;
data correl(type=corr);
_type_='CORR';
input _name_ $ x1 x2 x3 x4 x5 x6;
cards;
x1 1.0 . . . . .
x2 .8636 1.0 . . . .
x3 .6792 .6738 1.0 . . .
x4 .5099 .5164 .3274 1.0 . .
x5 .6347 .5738 .4165 .6751 1.0 .
x6 .6396 .6232 .4614 .7377 .8384 1.0
;
proc factor
method=ml
nfactors=2
rotate=varimax;
run;
proc factor
method=ml
nfactors=2
rotate=equamax;
run;
proc factor
method=ml
nfactors=2
rotate=orthomax;
run;
proc=factor
method=ml
nfactors=2
rotate=quartimax;
run;
proc=factor
method=ml
nfactors=2
rotate=parsimax;
run;
```

EJEMPLO EVALUACION ENSEÑANZA-APRENDIZAJE.
PRELIS

```
1                386 - P R E L I S  1.20
0                      BY
0               KARL G JORESKOG AND DAG SORBOM
```

This program is published exclusively by

SCIENTIFIC SOFTWARE, Inc.
1525 East 53rd Street, Suite 906
Chicago, Illinois 60615, U.S.A.
(800)247-6113 or (312)684-4979


```
0THE FOLLOWING PRELIS CONTROL LINES HAVE BEEN READ :


EVALUACION EDUCATIVA FACULTAD DE MEDICINA;
DA NI=16
LA
V1 V2 V3 V4 V5 V6 V7 V8 V9 V10 V11 V12 V13 V14 V15 V16
RA FI=A:PRUEBA.DAT
OR ALL
OU MA=PM SM=A:AWA.PML SA=A:AWA.ACP PA

0W_A_R_N_I_N_G : Sample size too small. Asymptotic variances and covariances
          will not be computed. ERROR CODE 215.
0        THE PROBLEM USED    118608 BYTES (= 45.2% OF AVAILABLE WORKSPACE)
```

LISREL BARTHOLOMEW.

```
1                386 - L I S R E L  7.20
0                      BY
0              KARL G JORESKOG AND DAG SORBOM


            This program is published exclusively by

                SCIENTIFIC SOFTWARE, Inc.
              1525 East 53rd Street, Suite 906
              Chicago, Illinois 60615, U.S.A.
              (800)247-6113 or (312)684-4979

   Copyright by Scientific Software, Inc. (a Michigan corporation), 1981-91.
  Partial copyrights by MicroWay Corp., 1988-91 and Phar Lap, Inc., 1986-91.
                    All rights reserved.
0THE FOLLOWING LISREL CONTROL LINES HAVE BEEN READ :


REPLICACION DEL ARTICULO DE ROTACION DE BARTHOLOMEW
DA NI=6 NO=100 MA=KM
LA
V1 V2 V3 V4 V5 V6
KM SY
1.0000
 .8636 1.0000
 .6792  .6738 1.000
 .5099  .5164  .3274 1.0000
 .6347  .5738  .4165  .6751 1.0000
 .6396  .6232  .4614  .7377  .8384 1.0000
MO NX=6 NK=2  PH=ST
LK
FACTOR1 FACTOR2
FR LX(1,1) LX(2,1) LX(3,1) LX(4,2) LX(5,2) LX(6,2)

OU ALL
```

BIBLIOGRAFIA

Bartholomew David J., (1987) *Latent variable models and Factor Analysis.,* Griffin's Statistical Monographs and Courses No. 40 Ed.. Alan Stuart., Oxford University Press, New York.

Bartholomew David J., (1980) *Factor Analysis for categorical data.,*J. R. Statistics Soc. 42 No. 3

Bartholomew David J., (1985) *The foundations of factor analysis.,* Biometrika 71,2.

Bartholomew David J., (1985) *Foundations of factor analysis: some practical implications.,*British Journal of Mathematical and Statistical Psychology, 38, 1-10.

Bernstein H. Ira., (1988) *Applied Multivariate Analysis.,*Springer-Verlag, New York.

Cattell B. Raymond.,(1987) *Intelligence: its Structure, Growth and Action.,* New York, Harper.

Cattell B. Raymond and Nesselroade., (1988) *Prospectives on individual diferences.,*Handbook of Multivariate Experimental Psychology., 2a. Ed. Cap. 4

Chistoffersson Anders.,(1975) *Factor analysis of dichotomized variables.,* Psychometrika, Vol. 40, No. 1, March.

Cronbach, L.., and Meehl, P., (1955) *Construct Validity in Psychological Test.,* Psychological Bulletin No. 52, pp 281-302.

Choynowski. M., (1988) *Ongoing Research: Aggressiveness and Personality.,* Person. Indiv. Diff. Vol. 9, No 1 197-198

Díaz Loving R. y Col., (1989) *.,Desarrollo y Análisis psicométrico de una medida multidimensional de celos.,* Revista Mexicana de Psicología, Jul-Dic Vol 6(2) 111-119

Dillon William R y Goldstein Matthew, (1984) *Multivariate Analysis.,* Wiley Sons.

Flores Rojas Ma. de la L. y Cuevas Renaud C, (1995)., *Estrategias de estudio en matemáticas.,*Coordinación de Matemáticas y Estadística, Fac. de Psicología, UNAM. Documento no publicado.

Guilford, Benjamin F., (1984) *Estadística aplicada a la Psicología y la Educación.,*México Mc Graw-Hill. Primera edición en español de la sexta edición en inglés *Fundamental Statistics in Psychology and Education* (1978).

Hair Joseph F. Jr. , Anderson E. Rolph, Tatham L. Ronald, Black C. William(1992) *Multivariate data Analysis.,*. Maxwell Macmillan.

Hope K., (1982) *Manual Práctico de Estadística Avanzada.,* México Trillas. Segunda reimpresión en español de la versión en inglés *The elementary Statistics (1967).* Pergamon Press Ltd, Londres, Inglaterra.

Jöreskog, K.G. (1963) *Statistical Estimation in Factor Analysis.,* Stockholm: Almqvist & Wiksell.

Jöreskog, K.G. y Sörbom.,LISREL Versión 7 (1988).,*A guide to the Program and Applications.,* SPSS.

Jöreskog, K.G. and Sörbon Dag.,(1979) *Advances in Factor Analysis and Structural Equation Models.,* Ed. Jay Magidson, Abt. Books Cambridge, Massachusetts.

Hernández, Cid Rubén., (1987) *Historie et Préhistorie de L'Analyse des Donneés.,* MATHESIS Vol. III No. 4 Noviembre de 1987., Departamento de Matemáticas, Facultad de Ciencias, UNAM.

Kaplan, David., (1991) *The behaviour of three weighted least squares estimators fo structured mean analysis with non-normal Likert variables.,* British Journal of Mathematical and Statistical Psychology 44, pp. 333-346.

Kendall, Maurice., (1980) *Multivariate Analysis.,* Macmillan Publishing Co. Inc., New York.

Kendall, Maurice and Stuart Alan., (1961) *The Advances Theory of Statistics.,* Volumen 2 Inference and Relationship., Charles Griffin & Company Limited, London.

Kerlinger, Fred N., (1981) *Enfoque Conceptual de la Investigación del Comportamiento.,* Ed. Interamericana México. Primera edición en español de la versión en inglés *Behavioral Research:a conceptual approach* (1979).

Krzanowski W.J. and Marriott FHC.,(1995) *Multivariate Analysis.* Part I Distributions, Ordination and Inference. Edward Arnold.

Krzanowski W.J. and Marriott FHC.,(1995) *Multivariate Analysis.* Part II. Edward Arnold.

Krzanowski W.J. (1988) *Principles of Multivariate Analysis.,* Oxford University.

Loehlin John C., (1992).,*Latent variable models and introduction to factor, path, and structural analysis.,* Editorial Lawrence Eribaum Associates.

Méndez Ramírez I., (1993) ., *El Paradigma Cuantitativo Vs. el Cualitativo en la investigación.,* Comunicaciones Técnicas, Monografias No. 110: Serie Azul, IIMAS, UNAM.

Mulaik, Stanley A., (1986) *Factor Analysis and Psychometrika:major developments.*,Psychometrika Vol. 51 No. 1, 23-33

Mulaik, Stanley A., (1972) *The foundations of Factor Analysis.*, Mc Graw-Hill Book Company.

Muthen B. and Kaplan D. (1992)., *A comparasion of some methodologies for the factor analysis of non-normal Likert variables: A note on the size of the model.*, British Journal of Mathematical and Statistical Psychology 45, 19-30.

Nordby J. Vernon y Hall S. Calvin., (1993) *Vida y conceptos de los Psicólogos más importantes.*, Trillas, p. 33 México. Primera edición en español de la versión en inglés de 1990.

Onkin-Jae and Mueller Charles W., (1982) *Introduction to factor analysis.*,Series Quantitative Applications in the Social Sciences Núm. 13 Sage University Paper.

Padua Jorge (1979).,*Técnicas de investigación aplicadas a las Ciencias Sociales.*, El Colegio de México. Fondo de Cultura Económica.

Peters, C. Charles.,*El análisis factorial múltiple.*,Cap. III .Documento inédito.

Potthast Margaret J. (1993)., *Confirmatory factor analysis of ordered categorical variables with large models.*, British Journal of Mathematical and Statistical Psychology 46 pp. 273-286.

Reyment R. and Jöreskog, K, G., (1993) Cambridge University Press.

Spearman Charles.,(1927) *The Abilities of Man.*,New York: Macmillan.

Spearman Charles.,(1904).,*General intelligence, objectively determined and measured.*,American Journal Psychology., **15**

Thurstone L. L. , Mann W. , Gurvitz M. S., Peters C, Brown F.J., Kinder F., Drake R.M.,G. Arthur y Berdie R.F., (1967) *Las capacidades básicas. La medición de la inteligencia, la aptitud y el interés.*, . Paidos., Cap. IV. Primera edición en español (1967) de la versión en inglés *Encyclopedia of Psychology.*(1947). Philosophical library, New York.

Thurstone L. L. , (1932, 1933) *The theory of multiple factors by L.L. Thustone.*, The University of Chicago. De. Edwards Brothers, Inc and Arbor Michigan.

Thurstone L. L. , (1969) *Multiple Factor Analysis.A development & expansion of the vectors of mind.* The University of Chicago Press.

Waller Niels G., (1993) *Seven Confirmatory Factor Analysis Programs: EQS, EZPATH, LINCS, LISCOMP, LISREL/, SIMPLIS, and CALIS.*, Vol. 17 No. 1, March 1993. Applied Psychological Measurement.

MANUALES

SAS (1991), *Language Guide for Personal Computers.*, Versión 6.03 .

SAS (1991), *Introductory Guide for Personal Computers.*, Versión 6.03.

LISREL (1989), **Jöreskog and Sörbom -SPSS** *A Guide to the Program and Applications,* 2a. Edición.

SPSS (1991)., **Marija J. Norusis.,** *SPSS for Windows,* Versión 5.0

PRELIS (1988)., **Jöreskog and Sörbom**., Scientifc Software, Incorporated.