Qüestiió - V. 7, n.º 1 (març 1983) pp. 359-383

ANÁLISIS DE LA CAUSALIDAD Y PLANTEAMIENTO LISREL A PARTIR DE LOS MODELOS DE MEDIDA

J. M. BATISTA*, C. M. CUADRAS**

* UNIVERSIDAD POLITÉCNICA DE BARCELONA

** UNIVERSIDAD DE BARCELONA

Partiendo del desarrollo de los modelos de medida del Análisis Estadístico multivariable, se presenta el modelo general LISREL, que permite analizar cualquier sistema estocástico de ecuaciones de estructura lineal, sin autocorrelación de los errores. Se describe brevemente la metodología exploratoria utilizada en Análisis Factorial, y se insiste en la adecuación del modelado confirmatorio en la validación de teorías. Se quiere establecer, en especial, la idoneidad del planteamiento LISREL como método de análisis de la causalidad en aquellos dominios en los cuales la experimentación resulte inapropiada o inviable.

CAUSAL ANALYSIS AND LISREL APPROACH FROM MEASUREMENT MODELS

Key words: Factor analysis, covariance structure analysis, path analysis, linear structure relationships, regression models.

1. INTRODUCCION.

Numerosas técnicas de Análisis Multivariable consisten en el estudio e interpretación de la información que proporcionan ciertas matrices, tales como: matrices de covarianza, de correlaciones, de similaridades, de disimilaridades, de dispersión entre grupos y dentro grupos, etc. Especial mención merecen aquellas técnicas directamente relacionadas con el Análisis Factorial, cuya finalidad es analizar la matriz de covarianzas, estudiando y verificando su estructura con el fin de poner de manifiesto la dimensionalidad latente de las variables observables originales. Desde los primeros modelos simples de AF, pasando por los modelos multifactoriales exploratorios y confirmatorios, se ha llegado en la actualidad a modelos generales para expresar la covarianza capaces de describir una gran variedad de modelos, de aplicación pluridisciplinar (Biología, Psicología, Economía, etc.) y que son el fundamento de los métodos mas recientes de análisis causal en la llamada investigación no experimental.

En el presente trabajo se utilizan las notaciones más aceptadas en la literatura sobre Análisis Multivariante, como son:

$x = (x_1, \ldots, x_p)'$	vector de p variables observables,
$\Sigma = E(x.x')$	matriz p×p de varianzas-covarianzas de las variables observables.
Λ	matriz factorial,
$\xi = (\xi_1, \ldots, \xi_r)'$	vector de r variables latentes,
$\delta = (\delta_1, \ldots, \delta_p)'$	vector de p factores únicos,
$E(\xi.\delta') = 0$	indicación de que las variables ξ están incorrelacionadas con las variables δ,
μ	vector de medias de las variables; normalmente se supone $\mu = 0$,
tr(A)	traza de la matriz A.

Asimismo utilizaremos abreviaturas para designar diferentes técnicas multivariantes:

ACP : análisis de componentes principales

AF : análisis factorial

AEC : análisis de estructuras covariantes

LISREL : "linear structural relationships"

J.M. Batista - Cátedra de Estadística Teórica y Aplicada de la E.T.S.E.I.B. de la Universitat Politécnica de Barcelona - Av. Diagonal, 647 - Barcelona

C.M. Cuadras - Dep. Bioestadística de la Fac.de Biologia de la Universitat de Barcelona - Av. Diagonal, 637 - Barcelona

- Article rebut el Juny del 1983.

ACOVS : programa para el análisis de estructuras covariantes

COFFAM : programa para el análisis factorial confirmatorio

2. ANALISIS DE COMPONENTES PRINCIPALES Y EL MODELO DEL ANALISIS FACTORIAL:

El primer método para analizar la dimensionalidad latente y significativa de p variables es el Análisis de Componentes principales (ACP), que consiste en transformar linealmente las variables en un nuevo conjunto de variables-componentes, incorrelacionadas entre sí y de varianza decreciente. La relación entre variables (x) y componentes (y) es

$$y = Vx \qquad (1)$$

siendo

$$V'V = I \qquad (2)$$

$$\Sigma = V\ D\ V' \qquad (3)$$

$$D = \mathrm{diag}\ (\lambda_1, \ldots, \lambda_p) \text{ con } \lambda_i \geq \lambda_{i+1}$$

La matriz $V(p \times p)$ contiene los p vectores propios de Σ. Los elementos de la diagonal de D son los valores propios. Por las propiedades que tienen las componentes principales (máxima variabilidad explicada, incorrelación mutua), el ACP se ha utilizado en Biología para establecer los conceptos de tamaño y forma de los individuos de una especie /1/ y para interpretar la noción de crecimiento /2/.

El ACP sugerido primeramente por K. Pearson /3/ y desarrollado por Hottelling /4/, /5/, se utiliza también como método de representación de datos, tomando como base una distancia euclídea (y entonces se diagonaliza Σ) ó una distancia euclídea normalizada (y entonces se diagonaliza R), en forma no muy distinta a la inicialmente presentada por K. Pearson, y que es el fundamento de las recientes técnicas del Análisis Factorial de Correspondencias /6/. En este sentido, el ACP ha sido aplicado: a la representación de especies, comunidades bentónicas, etc. /7/; a la representación de municipios en estudios de comportamiento electoral /8/, /9/; a la clasificación de los paises de la OCDE según dieciocho variables socioeconómicas /10/; a la definición de zonas homogéneas y estudio de la jerarquía urbana del sistema de ciudades de Cataluña /11/.

Las aplicaciones del AF a la psocología carecen, en general, de interés, pues desde que Ch. Spearmann /12/ avanzó la hipótesis de que todas las medidas intelectivas podían considerarse mediciones de un factor general "g", común a todas ellas, y de otro factor específico de cada medida, la tradición psicométrica ha modelado las variables observables $x(p\times 1)$ distinguiendo dos fuentes de variabilidad: una que origina la varianza común (comunalidad), y otra que da origen a la varianza única (unicidad). El AF es una extensión del ACP que postula un modelo de factores comunes, $\xi(k \times 1)$ y únicos $\delta(p \times 1)$. El modelo general es

$$x = \Lambda\xi + \delta \qquad (4)$$

verificando

$E(\xi.\delta') = 0$; $E(\delta.\delta') = \theta_\delta$, matriz diagonal,

es decir, existe incorrelación entre ambas fuentes de variabilidad (común y única), y los p factores únicos están mutuamente incorrelacionados, es decir, se han incluido todos los factores comunes en ξ y además se supone que los errores de medida están incorrelacionados. En estas condiciones se verifica:

$$E(x.x') = \Sigma = \Lambda\Phi\Lambda' + \theta_\delta \qquad (5)$$

donde $\Phi = E(\xi.\xi')$, es la matriz de covarianza de los factores comunes, que pueden ser ortogonales u oblicuos. θ_δ incluye la varianza específica y la del error de medida.

El AF en su sentido exploratorio (descubrir e interpretar la dimensionalidad latente) comprende las fases de selección de las variables, obtención de Σ, factorización de Σ y rotación de los factores obtenidos. El método del Factor Principal, el AF Canónico y el método Alfa, son los métodos más relevantes del AF exploratorio que permiten obtener soluciones factoriales directas. El número y la rotación de factores comunes se basa en ciertos principios de parquedad, cuya expresión analítica más aceptada es el criterio Varimax /13/. El enfoque exploratorio del AF se inicia en Spearman, tiene en Thurstone /14/ su principal defensor, y culmina con la obra de Harmann /15/. Véase también /16/.

2.1. INVARIANCIA POR CAMBIOS DE ESCALA.

Es conveniente decidir a priori la matriz -- (de correlaciones o de covarianzas) de la -- cual debe partirse para obtener la solución factorial. La solución obtenida en el ACP es distinta, y no se pasa de una a otra por un simple cambio de escala.

La invariancia de la solución factorial por cambios de escala es una propiedad que ha sido resaltada por Kaiser y Caffrey /17/. Los métodos Alfa y Canónico tienen esta propiedad, pero no el del Factor Principal. Swaminatham y Algina /18/ realizan un estudio -- preciso del concepto de invariancia de los - factores y concluyen que esta no depende tanto del método de estimación de la matriz factorial, como de las restricciones impuestas a los parámetros.

2.2. INTERPRETACION Y SIGNIFICADO DE COMPONENTES Y FACTORES.

Ya que el ACP y el AF intentan sustituir un conjunto de variables observables (x) por un reducido número de componentes o factores, - es interesante discutir las diferencias y -- ventajas entre uno y otro. En el ACP de (1) se tiene que x=Uy, donde U es invertible, y entonces

$$x = Uy = Uy_1 + Uy_2 \cong Uy_1 \qquad (6)$$

donde y_1 representan las primeras componentes principales consideradas. Las componentes pueden expresarse como función de las variables observables, luego son variables medibles y bien definidas. Pero la aproxima -- ción (6) es menos natural que la considerada en AF:

$$x = \Lambda\xi + \delta \cong \Lambda\xi \qquad (7)$$

donde se admite que x es una expresión de -- los factores salvo un error que se expresa - en forma de unicidad. Sin embargo, nos encontramos entonces con un problema. Los factores comunes no son variables directamente medibles, sino de naturaleza hipotética, violando por tanto el principio experimental -- que considera como válidos sólo aquellos entes que son observables. Cuando se dice que la variabilidad biométrica de los individuos de una especie se explica solamente por una componente que expresa "tamaño", estamos reduciendo las variables a una sola que puede ser medida y observada. Pero la afirmación - de que ciertos tests de personalidad están - dominados por dos dimensiones (introversión-extraversión y neuroticismo-estabilidad), -- significa admitir la existencia de unas variables que no podemos observar. En general, las aplicaciones que se han hecho del AF a - las ciencias de la conducta tienen cierto -- transfondo especulativo, donde unas causas, cuyo significado no está suficientemente claro, quieren justificarse y explicarse a través del modelo (7). Véase /19/.

3. MODELOS PARA LA CONFIRMACION DE TEORIAS.

El problema de decidir si un conjunto de datos se ajustan a un determinado modelo, confirmando una teoría previamente formulada -- (de naturaleza genética, psicológica, económica, etc.), es uno de los más generalizados en Estadística. La prueba ji-cuadrado para - la comparación de frecuencias esperadas con frecuencias observadas es un ejemplo clásico. Este punto de vista ha afectado también al - AF. Si el enfoque exploratorio pretende descubrir la dimensionalidad latente sin fijar una estructura previa (salvo ciertos criterios de parquedad), el enfoque confirmatorio del AF postula, parcial o totalmente, un modelo estructural para los factores. Varias - son las situaciones que justifican el AF confirmatorio:

1) la estructura de los datos y las variables se ajusta a una teoría concreta;

2) resultados de estudios exploratorios previos sugieren como debe ser el modelo factorial y la naturaleza de los factores;

3) la estructura de los datos obedece a un - diseño clasificatorio de items según rasgos objetivos de contenido y/o formato.

3.1. ANALISIS DE ESTRUCTURAS COVARIANTES.

En su formulación general, la confirmación - de un determinado modelo puede plantearse estudiando hasta que punto la matriz de covarianzas de las variables observables obedece a una cierta estructura aditiva-multiplicativa, que depende del modelo planteado. El Análisis de Estructuras Covariantes (AEC) es el método general que permite resolver este problema. Es un método que describe una gran variedad de modelos, que permiten detectar y confirmar fuentes de variación entre las varianzas y covarianzas de las variables observadas.

Las ideas previas del AEC se encuentran en los trabajos de Bock /20/ y Bock y Bargmann /21/, pero es la versión de Jöreskog /22/ la que se acepta actualmente por su mayor generalidad. Véase también /23/.

Sea $y = (y_1, \ldots, y_p)'$ un conjunto de p variables aleatorias. El modelo general es:

$$y = \mu + \Lambda\Gamma\xi + \Lambda\zeta + \varepsilon \tag{8}$$

$$\Sigma = \Lambda\ (\Gamma\Phi\Gamma' + \Psi^2)\ \Lambda' + \theta_\varepsilon \tag{9}$$

siendo $\Lambda(pxq)$, $\Gamma(qxr)$, $\Phi(rxr)$ matriz simétrica, $\Psi^2 = \mathrm{diag}\ (\psi_1^2, \ldots, \psi_q^2)$ $\theta_\varepsilon = \mathrm{diag}\ (\theta_{\varepsilon 1}, \ldots, \theta_{\varepsilon p})$. Los vectores aleatorios

$$\xi = (\xi_1, \ldots, \xi_r) \quad \zeta = (\zeta_1, \ldots, \zeta_p)' \quad \varepsilon = (\varepsilon_1, \ldots, \varepsilon_p)'$$

se suponen incorrelacionados, centrados y con matrices de covarianzas

$$E(\xi.\xi') = \Phi \quad E(\zeta.\zeta') = \Psi^2 \quad E(\varepsilon.\varepsilon') = \theta_\varepsilon$$

El número de variables p es conocido, pero los valores de q y r dependen de argumentos teóricos. Los parámetros que intervienen en $\pi = (\Lambda, \Gamma, \Phi, \Psi, \theta_\varepsilon)$ pueden ser de tres tipos:

a) fijos, es decir, con valores asignados conocidos,

b) restringidos, es decir, desconocidos pero algunos de ellos iguales,

c) libres, es decir, desconocidos y no restringidos.

Los dos aspectos que deben resolverse en AEC son:

1) Estimación de los parámetros, teniendo en cuenta las posibles restricciones impuestas,

2) Confirmación del modelo propuesto mediante test de hipótesis estructurales.

La resolución de los mismos es bastante complicada. La estimación puede llevarse a término según el criterio de los mínimos cuadrados generalizados o por el método de la máxima verosimilitud. El test de hipótesis se describe mediante un estadístico ji-cuadrado. Véase/22/ y /24/.

El AEC sintetiza numerosas técnicas multivariantes, cuyo principal campo de aplicación es la psicología. Aunque cada modelo particular es preferible describirlo independientemente, tiene la ventaja práctica de que puede ser programado globalmente. El programa ACOVS /25/ donde se encuentra implementado este modelo AEC se considera absoleto en la actualidad, pues como se verá en el apartado 5.2 es una particular especificación del modelo LISREL, que fué divulgado en el libro de Goldberger y Duncan /26/.

3.2 ESTRUCTURA DE PENROSE.

La estructura más simple de la covarianza es la propuesta por Penrose /27/ en problemas de morfometría. Supone que todas las variables son equicorrelacionadas, con lo cual Σ depende solamente de dos parámetros: σ^2 (varianza) y ρ (correlación). Existe entonces una única componente principal (hay isometría a partir de la primera). En aplicaciones biométricas esta estructura significa que los individuos presentan crecimiento isométrico respecto a las medidas consideradas, y la variabilidad depende solamente de la componente "tamaño".

3.3 ANALISIS FACTORIAL CONFIRMATORIO.

El enfoque confirmatorio del AF tiene su origen en el trabajo de Jöreskog /28/, aun cuando la etapa de especificación en la modelización se origina propiamente en los modelos de ecuación estructural con la comprobación de teorías, en la línea de la investigación no experimental (Véase /29/).

El modelo del AF confirmatorio y la estructura de Σ que de él se deriva es

$$x = \mu + \Lambda\xi + \delta \tag{10}$$

$$\Sigma = \Lambda\Phi\Lambda' + \Psi^2 \tag{11}$$

donde ξ son los factores comunes, δ los factores únicos, Λ la matriz del modelo factorial, y Φ la matriz de intercorrelaciones entre los factores comunes. Se trata de un caso particular del AEC donde

$$\Lambda = I \qquad \theta_\varepsilon = 0$$

El planteo dado en /28/ tiene ventajas sobre otros propuestos con anterioridad (Howe /30/, Anderson y Rubin /31/, Lawley y Jöreskog /32/) pues los elementos fijos de las matrices no tienen que ser necesariamente nulos, no es preciso que las restricciones sean las suficientes para que la solución sea única, etc.

Los tests de hipótesis en AF confirmatorio se realizan sobre:

1) El número r de factores comunes;

2) La naturaleza de los factores, que pueden ser ortogonales, oblícuos y mixtos.

3) Las saturaciones, Λ, del modelo factorial que pueden ser libres, restringidas o fijas.

4) Las restricciones impuestas al vector de medias μ .

Joreskog /28/, /33/, Mulaik /34/, Lawley /35/ Cuadras /16/ exponen diferentes aplicaciones a la psicología. Análogamente a lo que sucede en el programa ACOVS, el programa COFFAM /36/, en el que se implementó este modelo, tampoco está vigente en la actualidad.

3.4 MODELOS DE LA TEORIA CLASICA DEL TEST.

El AEC incluye como casos particulares ciertos modelos para conjuntos de tests de puntuación verdadera, desarrollados por Lord y Novick /37/ y formulados posteriormente por Jöreskog /38/.

Las ciencias llamadas exactas solventan usualmente mediante la"replicación" aquellos errores de medida en que hubiesen incurrido los instrumentos de que se sirven para medir. En Psicología y ciencias sociales en general, la repetibilidad en la experimentación o es impracticable, o bien se distorsiona debido a efectos de memoria, tedio y cansancio. La solución adoptada en estas disciplinas para evitar en lo posible errores de medida, se basa en la utilización de tests congenéricos, tau-equivalentes y paralelos, que se supone miden la misma variable de distintas maneras. Según estos planteamientos la puntuación observada en la i-ésima variable del test consta de dos componentes: una sistemática que representa la verdadera puntuación (τ_i); y otra aleatoria o error de medida (e_i), donde tanto τ y e, como los errores entre sí, están incorrelacionados:

$$x_i = \tau_i + e_i \tag{12}$$

Una indicación de la calidad del procedimiento de medida es el llamado coeficiente de fiabilidad que se define como: $E^2(x_i . \tau_i)$.

Se dice entonces que dos medidas son congenéricas si la correlación entre todo par de verdaderas medidas es igual a la unidad:

$$\text{corr } (\tau_i, \tau_j) = 1 \qquad i,j = 1,\ldots,p \tag{13}$$

Si suponemos que cada τ_i tiene media cero y varianza unidad, el modelo que describe el conjunto de tests congenéricos es

$$x = \mu + \lambda\tau + e \tag{14}$$

donde $\lambda=(\lambda_1, \ldots, \lambda_p)'$ es el vector de coeficientes de regresión. En las condiciones impuestas, la matriz de covarianzas de las variables observables es:

$$\Sigma = \lambda\lambda' + \theta \tag{15}$$

donde $\theta = \text{diag } (\theta_1, \ldots, \theta_p)$, y $\theta_i = \text{var } (e_i)$. Es un caso especial de (9) siendo

$$\Gamma = \lambda \qquad \Lambda = \Phi = I \qquad \Psi = 0$$

Los tests paralelos son casos particulares en los que

$$\lambda_1 = \ldots = \lambda_p \qquad \theta_1 = \ldots = \theta_p$$

mientras que en los tau-equivalentes se admite igual varianza para las verdaderas medidas, pero posiblemente diferentes varianzas en los errores. Esta última situación es propia de los instrumentos de medida que midiendo la misma variable,con igual unidad de medida, no tienen la misma precisión.

En las ciencias sociales y de la conducta re sulta práctica habitual tomar como medida de fiabilidad el cuadrado del coeficiente de correlación entre medidas repetidas, no obstante, como se desprende de lo anterior, esto sólo es correcto cuando el instrumento de medida es paralelo o tau-equivalente, ya que unicamente en estos casos $E^2(x.x')$ coincidirá con λ^2.

A pesar de la similitud de las expresiones (10) y (14), deben advertirse dos diferencias esenciales entre el modelo de AF y los de la teoría clásica del test:

1ª. En (14) el término de error e, no incluye la componente específica que se le suponía a δ en (10). Así la sobreestimación de la varianza del error en que se incurre al tomar (10) en lugar de (14) redunda en calcular erróneamente el índice de fiabilidad anterior.

2ª. Si el modelo (14) se generaliza a k conjuntos múltiples de medidas congenéricas, con $m_1,\ldots,m_k$ tests respectivamente, la matriz de coeficientes Λ, a diferencia de (10), será una matriz diagonal de bloques que expresará la menor complejidad de estos tests con respecto a las variables observables del modelo de AF.

Generalizaciones de estos tests pueden encontrarse en Kristof /39/ y Alwin y Jackson /40/. Recientemente y en diversas disciplinas, se han llevado a cabo aplicaciones de la utilización de la teoría clásica del test para evaluar la validez de algunos instrumentos de medida habitualmente considerados como adecuados. Véase Jansen /41/ y Saris /42/.

3.5 ESTIMACION DE COMPONENTES DE LA VARIANZA Y COVARIANZA.

Bock /20/, y Bock y Bargmann /21/, proponen un método para el análisis de las componentes de la varianza en diseños preestablecidos, por asignación de items a categorías de una clasificación factorial o jerárquica, según características objetivas de contenido o formato. En estos diseños se considera un factor aleatorio (sujetos), y los otros factores como fijos.

El modelo en notación vectorial es

$$y = \mu + \eta + e \qquad (16)$$

con $\eta = \Gamma\xi$

donde η es el vector de verdaderas puntuaciones que representa no sólo la variable objeto de estudio, sino también las diversas componentes del diseño; e es el vector de los errores de medida; ξ incluye las p componentes aditivas de clasificación, y Γ es una matriz cuyos elementos son en general conocidos (matriz diseño). La estructura de la covarianza es

$$\sum = \Gamma\Phi\Gamma' + \Psi' \qquad (17)$$

siendo Φ la matriz diagonal de las varianzas de las componentes o modalidades de clasificación, y Ψ^2 la matriz diagonal de las varianzas residuales. La estructura (17) es un caso particular del AEC en donde Λ=I, θ=0. La diferencia con (9) estriba en que la matriz de saturaciones Γ debe ser estimada, mientras que en (17), Γ es una matriz, en general, conocida. Se trata de estimar las componentes de la varianza y covarianza, y decidir sobre la validez del modelo (16).

Análogo planteamiento al de estos autores es el de Cronbach y otros /43/. Los modelos posteriores de Wiley y otros /44/, Mulaik /45/, Jöreskog /46/, y Mallenberg y otros /47/, representan planteamientos más generales al relajar algunas de las restricciones impuestas en (16) y (17).

Dekkin y Raadsheer /48/, construyen un instrumento para medir la ansiedad (ε_1) de los niños según un diseño de dos factores: uno situacional (ε_2 a ε_6); y otro (ε_7 a ε_9) conductual, cuya matriz de diseño Γ, es:

	ξ_1	ξ_2	ξ_3	ξ_4	ξ_5	ξ_6	ξ_7	ξ_8	ξ_9
y_1	x	x	0	0	0	0	x	0	0
y_2	x	x	0	0	0	0	0	x	0
y_3	x	x	0	0	0	0	0	0	x
y_4	x	0	x	0	0	0	x	0	0
y_5	x	0	x	0	0	0	0	x	0
y_6	x	0	x	0	0	0	0	0	x
y_7	x	0	0	x	0	0	x	0	0
y_8	x	0	0	x	0	0	0	x	0
y_9	x	0	0	x	0	0	0	0	x
y_{10}	x	0	0	0	x	0	x	0	0
y_{11}	x	0	0	0	x	0	0	x	0
y_{12}	x	0	0	0	x	0	0	0	x
y_{13}	x	0	0	0	0	x	x	0	0
y_{14}	x	0	0	0	0	x	0	x	0
y_{15}	x	0	0	0	0	x	0	0	x

en donde las x, representarían los elementos considerados no nulos que en general serán - unos de la matriz. Véase en Saris /49/ la -- comparación de los modelos propuestos por -- Bock /20/, Wiley y otros /44/, Mallenberg y otros /47/, para los datos del estudio de - Dekkin y Raadsheer /48/

3.6 MATRICES MULTIRASGO-MULTIMETODO

Los tests congenéricos planteados en el apartado 3.4, fueron propuestos por Campbell y Fiske /50/ para evaluar la validez y adecuación de n métodos en la medida de m rasgos. Este tipo de diseño factorial, considera que la verdadera puntuación (τ) no sólo refleja la variable teórica o rasgo que quiere medirse, ξ_j, sino también una componente específica, ξ_k, del método de medición utilizado.

El modelo considera a la i-ésima variable -- observable:

$$x_i = \tau_i + e_i \qquad (18)$$
$$\tau_i = \xi_j + \xi_k$$

La matriz Σ , donde se correlacionan las n.m medidas, se la conoce como la matriz multi-- rasgo-multimétodo y es un caso particular de la ecuación (9) de AEC, así

$$\Sigma = \Lambda\Phi\Lambda' + \theta \qquad (19)$$

donde Φ es una matriz diagonal de bloques debido a la incorrelación entre rasgos y métodos, θ es la matriz diagonal de las varian-- zas de los errores, y Λ tiene forma distinta según los métodos sean congenéricos o no respecto a la medida del rasgo.

En general la matriz Λ no es de rango completo, pudiéndose reparametrizar según indica - Jöreskog /51/ y analizar por procedimientos - de AF confirmatorio.

Aplicaciones de estas matrices pueden encontrarse en Werts y otros /52/, Jöreskog /51/, y Alwin /53/. Saris y otros /54/, estudian - el efecto que las distintas modalidades de - respuesta (categorías, C; líneas, L; números N) tienen en el modelo que mediante tres variables (identificación con el partido, ξ_1;- confianza con el candidato, ξ_2; preferencia por el candidato, ξ_3)describe la "tendencia - del voto", según el diagrama causal adjunto,

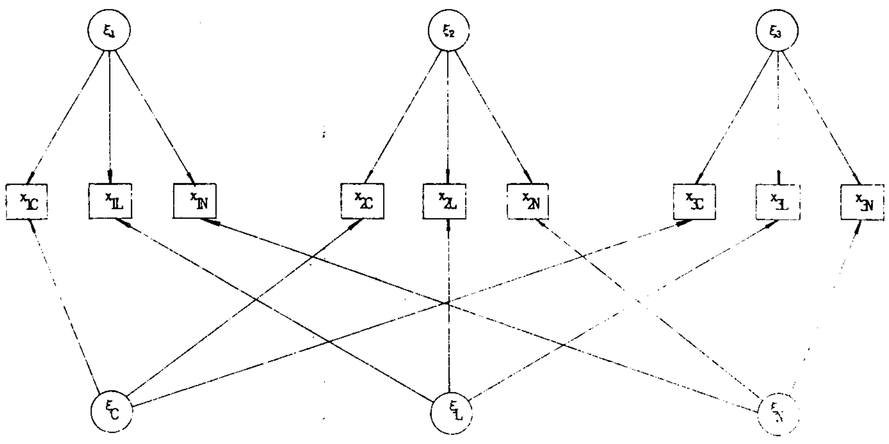

Fig. 1: Modelo de medida del que se sirven Saris y otros /54/. En el diagrama se ha omitido el proceso causal subyacente entre las variables de la "tendencia del voto" ξ_j.

3.7 MODELOS SIMPLEX Y CIRCUMPLEXOS.

Sea $x = (x_1,...,x_p)'$, y supongamos que existe una escala ordinal que permite definir un orden natural entre las variables $x_1,...,x_p$. - En Biología puede tratarse de una variable de crecimiento que se mide en períodos ordenados de tiempo. En Ecología, medidas de abundancia de especies encontradas en lozalizaciones diferentes pueden ordenarse de acuerdo con una variable de distancia geográfica. En Psicología, una serie de tests que supongan una dificultad progresiva implican su aplicación secuencial. Admitiremos, pues, en general, que existe un parámetro de escala, t, tal que

$$t_1<t_2<...<t_p$$

de modo que la información multidimensional - es

$$x = (x(t_1),...,x(t_p))'$$

Guttman /55/, estableció por primera vez la - estructura simplex para explicar las correlaciones entre tests ordenados, caracterizados por la gradación de estas correlaciones a medida que se alejan de la diagonal principal - de la matriz de correlaciones. Anderson /56/, formula la estructura de estos modelos bajo - la perspectiva de los procesos estocásticos - de Wiener y Markov, distinguiendo entre el modelo simplex del quasi-simplex según que el - error de medida de cada variable se ignore o sea añadido. Los modelos simplex de Wiener dependen de la unidad de medida de las variables, y sólo son aplicables cuando esta es la misma para todas ellas. Sin embargo, para los modelos simplex de Markov, la invariancia por cambios de escala de las variables permite -- trabajar con unidades de medida arbitrarias. La conexión entre las correlaciones y la escala t (siendo $t_i<t_j$) es la siguiente:

$$\rho_{ij} = \sqrt{t_i/t_j} \quad \text{(modelo simplex de Wiener)} \quad (20)$$

$$\rho_{ij} = \rho^{|t_i-t_j|} \quad \text{(modelo simplex de Markov)} \quad (21)$$

Para el modelo (21) la matriz de covarianzas es

$$\Sigma = DRD \quad (22)$$

donde R es la matriz de correlaciones y D es la matriz diagonal con las desviaciones tipo de las variables. Para el modelo simplex --- (20) es

$$\Sigma = TD_{t^*}T' \quad (23)$$

siendo $D_{t^*} = \text{diag}\,(t_1^*,...,t_p^*)$, $t_i^* = t_i-t_{i-1}$, y T la matriz

$$T = \begin{pmatrix} 1 & 0 & . & . & . & 0 \\ 1 & 1 & 0 & . & . & 0 \\ 1 & 1 & 1 & 0 & . & 0 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & 0 \\ 1 & 1 & . & . & . & 1 \end{pmatrix}$$

Los modelos quasi-simplex se obtendrían añadiendo una matriz θ a $\sum$.

Si las variables se asimilan a un orden circular en lugar de la ordenación lineal anterior, se habla de modelos circumplexos.

Ejemplos y aplicaciones de estos modelos se encuentran en Guttman /55/ y Jöreskog /51/.

3.8 MODELOS CAUSALES.

Finalmente el AEC permite tratar también los modelos del "Path analysis", técnica que introduce Sewall Wright /57/, /58/, y que consiste en:

1º) Representar mediante grafos el proceso causal teórico, atendiendo a ciertas convenciones intuitivas que hacen acordes estos gráficos con los sistemas de ecuaciones que describen el modelo causal.

2º) Descomponer las medidas de covariación en función de los parámetros del modelo (posteriormente Duncan /29/, formalizará esta técnica de análisis basándose en el álgebra de Esperanzas matemáticas), para estimar la magnitud de las relaciones lineales entre las variables, es decir, los parámetros del sistema de ecuaciones.

3º) Aislar la contribución de cada variable causal por separado, diferenciando efectos directos e indirectos de otras relaciones no causales, con lo que se obtiene información acerca de las hipótesis causales propuestas e incluso, se promueve la generación de otras nuevas.

Aunque las líneas en las que se han desarrollado los modelos causales han sido principalmente la econométrica y la sociométrica, Jöreskog /22/, muestra como el AEC permite resolver algunos de los problemas inherentes a las técnicas de regresión utilizadas en la estimación de los parámetros del modelo, especialmente, la inconsistencia de los estimadores frente a errores de medida en las variables endógenas, y la contrastación de la estructura causal en su conjunto. Estas últimas décadas, con la inclusión de variables no observables en los estudios de la causalidad, la inadecuación de las técnicas del "Path analysis" ha sido patente frente a: modelos con múltiples indicadores y múltiples causas (MIMIC models); modelos longitudinales con variables latentes; errores de medida correlacionados, etc. Jöreskog /51/, mediante el AEC proporciona soluciones a estos problemas.

Aplicaciones de estos modelos pueden encontrarse en Werts y Linn /59/, Hauser y Goldberger /60/, y Jöreskog y Goldberger /61/, entre otros.

4. LA INVESTIGACION EN LAS CIENCIAS HUMANAS Y SOCIALES.

Tradicionalmente el estudio de la causalidad ha puesto el énfasis en el diseño, que establece como recoger los datos de manera que se excluyan las posibles explicaciones alternativas. Mantener controladas todas las variables relevantes, excepto aquella cuyos efectos son objeto de estudio, es la base de la investigación experimental (Véase, Arnau /62/), aceptada hasta ahora como la más adecuada para comprobar hipótesis causales. En las investigaciones sociales y de la conducta, problemas éticos y de operatividad limitan las posibilidades de implementación del experimento clásico y de su posterior generalización desde el laboratorio; el control experimental resulta en general inasequible, debido a que, por un lado las variables causales no son manipulables por el investigador, y por otro lado, los individuos estudiados no sólo difieren en el aspecto para el cual se desea establecer el efecto causal; así los resultados observados pueden provenir igualmente (parcial o totalmente) de la influencia de otras variables. Todo ello ha conducido a que en las últimas décadas, la investigación en las ciencias humanas, sociales y biológicas (Psicología, Psicología social, Biología, Economía, etc.), recurra cada vez más a métodos de control no experimentales (estadísticos), que son aplicados una vez los datos han sido recogidos.

4.1 UTILIZACION DE MODELOS EN EL ANALISIS DE LA CAUSALIDAD.

Normalmente la inferencia científica se lleva a cabo de acuerdo con el paradigma hipoté tico-deductivo, según el cual;

1º) Se supone un modelo que "estructura" lo no observable,

2º) Se deducen consecuencias observables para el modelo propuesto

3º) Se realiza una investigación empírica con el objetivo de contrastar estas consecuencias esperadas con los datos recogidos. - Véase en la Fig. 2 el gráfico que Saris y Stronkhorst /63/ proponen para el proceso completo.

Así pues, toda inferencia y en particular la causal, supone rechazar o confirmar hipótesis según los datos recogidos. La interpretación de los datos, requiere un conjunto limitado - de supuestos acerca de como los datos han sido generados. Al número mínimo de asunciones que, como instrumento metodológico, son capaces de estructurar aquellos datos según una - cierta teoría, es lo que entendemos por modelo. Bentler /64/, afirma que los modelos causales cumplen una función de puente entre la teoría y la investigación aplicada.

El modelo causal completo que se presenta en el próximo apartado combina una parte estructural con un modelo de medida. La parte es-- tructural del modelo describe las supuestas relaciones entre las variables no observa--- bles que surgen cuando las variables medidas difieren de sus contrapartidas teóricas. Las variables no observables, representan a menudo constructos (a los que se asigna ciertos valores para tratarlos como variables latentes) cuyas implicaciones determinan las relaciones entre las variables observables en el sentido del AF (apart. 2). Así, la parte de medida del modelo proporcionará la ligazón - entre constructos o variables no medibles y las variables observables. El objetivo es verificar las afirmaciones causales hechas entre las variables del modelo a partir de las relaciones observadas en los indicadores.

Los modelos de medida han sido el objeto de los apartados precedentes. Consúltese la literatura econométrica y sociométrica apropiada para el estudio de los modelos estructurales.

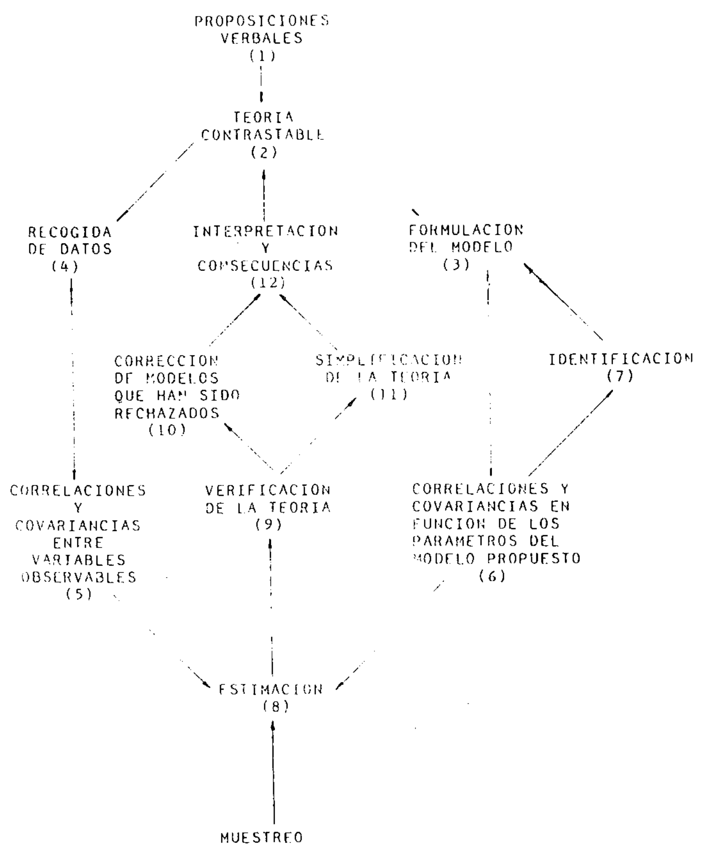

Fig.2 DIAGRAMA DEL PROCESO DE VALIDACION DE HIPOTESIS CAUSALES CON DATOS NO EXPERIMENTALES

(Saris y Stronkhorst, 1981)

5. MODELO GENERAL PARA ANALIZAR ECUACIONES DE ESTRUCTURA LINEAL. PLANTEAMIENTO LISREL.

5.1 INTRODUCCION.

En las últimas décadas, ha sido la covariación el único medio, no experimental, para revelar la causalidad. Los trabajos de Simon /65/, y la técnica de la correlación parcial de Simon y Blalock (ver Blalock /66/) estudiaban bajo qué condiciones la correlación no nula era evidencia suficiente para inferir relación causal. Concretamente estos últimos, utilizando la correlación parcial, demostraron que si en una teoría causal se especifican todas las variables relevantes, pueden deducirse ciertas conclusiones contrastables con respecto a las covariaciones entre variables. Posteriormente los modelos estructurales desarrollados por Boudon /67/, y Duncan /68/ basándose en el trabajo de S. Wright /57/, establecen la relación exacta entre efectos causales y medidas de covariación, estimando por ende dichos efectos.

Desafortunadamente, según Saris y Stronkhorst /63/, el salto de Simon-Blalock al "Path-Analysis" tuvo un indeseable efecto colateral al centrarse los investigadores en los procedimientos de estimación de los parámetros causales más que en la comprobación de teorías. No obstante, al enfrentarse las ciencias sociales (econômetras y sociómetras principalmente) con problemas concernientes al error de medida (en las variables), análogos a aquellos que los psicólogos trataban en los modelos de Análisis Factorial, aparecieron nuevos métodos que conjugaban el rigor en el "testing" con la eficiencia de la estimación. Este desarrollo fructifica en la conferencia interdisciplinaria que Goldberger convocaría en 1971, y de la cual surgió el planteamiento general: LINEAR STRUCTURAL RELATIONSHIPS. LISREL; o modelo JKW: Jöreskog /69/; Keesling /70/; Wiley /71/. Las conclusiones a las que llegaron en aquella conferencia, se recogen en el libro de Goldberger y Duncan /26/.

La identificación del sistema se debe a Wiley /71/; la estimación y verificación consiguientes son obra de Jöreskog /69/. La implementación de este sistema general por Jöreskog y Van Thillo /72/, se encuentra en el programa LISREL III; o en el LISREL IV y V de Jöreskog y Sörbom /73/, /74/.

5.2 FORMULACION DEL MODELO GENERAL.

El modelo causal que se considera supone:

1º) La existencia de relaciones lineales entre variables endógenas "η_i" $(i=1,\ldots,m)$ y predeterminadas "ξ_j" $(j=1,\ldots,n)$, que sin pérdida de generalidad se suponen centradas para evitar así una constante en la ecuación. El sistema más general (que será el no recursivo) de ecuaciones simultáneas, se expresa en forma matricial según

$$B\eta = \Gamma\,\xi + \zeta \tag{24}$$

La expresión (24) se conoce como la "ecuación estructural del sistema". Esta ecuación, en la que se supone un efecto aditivo del término de error ζ (mx1), no refleja simplemente un conjunto de ecuaciones lineales, sino que representa los mecanismos causales entre variables latentes, en los que reside la explicación de los efectos causales (B (mxm); Γ (mxn)) y la cantidad de varianza no explicada (Ψ).

De la exhaustividad de la teoría, los econômetras asumen que las variables de la ecuación (24) cumplen:

$E(\zeta) = 0$
$E(\xi\zeta') = 0$
B es no singular.

Habitualmente, el modelo de la ecuación (24) se expresa en la forma:

$$\eta = \beta\,\eta + \Gamma\,\xi + \zeta \tag{25}$$

esto es:

$$(I-\beta)\eta = \Gamma\,\xi + \zeta$$

Obviamente, la matriz que se nota en la ecuación (24) por B es $(I-\beta)$, y no se corresponde con la de efectos directos entre variables endógenas, sino que ésta sería $\beta=I-B$.

2º) Atendiendo a que en general η y ξ no serán exactamente medibles o bien no observables, se representará con y el vector que contiene a los indicadores de las variables endógenas, siendo x el que represente a los indicadores de las variables exógenas.

Si se asume de nuevo que indicadores (x,y) y constructos (ξ,η) están relacionados linealmente, se llega a formular, con la notación propia del AF, las dos "ecuaciones de medida" siguientes:

$$\begin{aligned} y &= \mu + \Lambda_y \ \eta + \varepsilon \\ x &= \nu + \Lambda_x \ \xi + \delta \end{aligned} \qquad (26)$$

En (26) las matrices paramétricas $\Lambda_x(qxn)$, $\Lambda_y(pxm)$, denotan que es posible la influencia de varios constructos en una misma variable (no es aconsejable, en general, la biyectividad entre constructos e indicadores: $\Lambda_x = \Lambda_y = I$). Estas matrices, junto con θ_ε y θ_δ, se utilizan para describir las propiedades de medida (validez y fiabilidad) de las variables observables, ya que el modelo de medida especifica el anclaje en términos reales de variables latentes y constructos hipotéticos.

La literatura psicométrica, principalmente la del AF, ha recogido los siguientes supuestos asociados a las ecuaciones de medida:

$$E\ (\eta\varepsilon') = E\ (\xi\ \delta') = 0$$
$$E\ (\varepsilon\delta') = 0$$
$$E\ (\varepsilon) = E\ (\delta) = 0$$

Debido a la combinación de las distintas ecuaciones en un modelo único, deben recogerse algunos supuestos adicionales.

i) incorrelación entre variables omitidas y errores de medida:

$$E\ (\zeta\varepsilon') = E\ (\zeta\delta') = 0$$

ii) incorrelación entre errores de medida y variables exógenas y endógenas:

$$E\ (\xi\varepsilon') = E\ (\eta\delta') = 0$$

Al conjugar los dos sistemas de ecuaciones: estructural y de medida, se revela el modelo general como la combinación de uno econométrico no recursivo y de dos psicométricos de Análisis Factorial, en los que los factores están relacionados linealmente (según la ecuación estructural). Así, los económetras abarcan el caso en que las variables no son medibles directamente y el modelo factorial se extiende al caso en que se pretende explicar la covariación entre factores, partiendo de las puntuaciones en otras variables manifiestas. El modelo de medida (26) puede escribirse de manera más compacta como:

$$z = \Lambda\ \Omega + e \qquad (27)$$

donde:

$$z = (y',x')'; \ \Omega = (\eta',\xi')'; \ e = (\varepsilon',\delta')';$$

$$\Lambda = \begin{bmatrix} \Lambda_y & 0 \\ 0 & \Lambda_x \end{bmatrix}$$

y puede interpretarse como un modelo restringido de AF (ver Jöreskog /27/), donde los factores de primer y segundo orden, satisfacen un sistema lineal de ecuaciones estructurales de la forma (24). Sin embargo, aquí no se exige que: $m<p$ y $n<q$ ó que θ_ε y θ_δ sean diagonales, como en el AF tradicional. Sólo se exige que la matriz de covarianzas de las variables observables Σ, sea no singular y que el modelo sea identificable (ver apart. 5.4).

Todas las teorías que se asuman con errores de medida y que incluyan relaciones lineales y estocásticas entre las variables, pueden formularse como casos especiales del sistema, determinando el número de variables explicativas y explicadas e indicando la correspondencia entre variables de la teoría y del modelo. Conviene también concretar las variables cuyo efecto se suponga nulo (u otro valor fijo) y cuáles no, es decir, especificar los parámetros fijos, restringidos y libres de las ocho matrices paramétricas siguientes:

B: El elemento β_{ij} de esta matriz indica el efecto directo de la j-ésima variable endógena en la i-ésima variable endógena, con el signo cambiado (ver ecuación (25)).

Γ: El elemento γ_{ij} de esta matriz denota el efecto directo de la j-ésima variable predeterminada en la i-ésima variable endógena.

Ψ: E $(\zeta\zeta')$: es la matriz de varianzas-covarianzas de los residuos, términos de error en las ecuaciones estructurales.

Φ: E $(\xi\xi')$: es la matriz de varianzas-covarianzas de las variables no observables predeterminadas.

Λ_x: El elemento λx_{ij}, de esta matriz, indica el efecto directo de la j-ésima variable predeterminada no observable en la i-ésima variable predeterminada observable.

$\theta_\delta = E(\delta\delta')$: es la matriz de covarianzas de los errores de las variables exógenas y predeterminadas de medida (indicadores).

Λ_y: El elemento λy_{ij}, de esta matriz indica el efecto directo de la j-ésima variable endógena no observable en la i-ésima variable endógena observable (indicador).

$\theta_\varepsilon = E(\varepsilon\varepsilon')$: es la matriz de varianzas-covarianzas de los errores de las variables endógenas de medida.

5.3 ESTRUCTURA DE LA MATRIZ DE VARIANZAS-COVARIANZAS DE LAS VARIABLES OBSERVABLES.

Si previamente se expresa la relación entre las variables no observables en su forma reducida, expresando cada variable endógena en función solo de las exógenas y predeterminadas puede llegarse, sin dificultad, a obtener la matriz C de varianzas-covarianzas de las variables latentes como función de los parámetros del modelo (Véase, para mayor información, Saris /49/):

$$C = \begin{bmatrix} C_{\eta\eta} & C_{\eta\xi} \\ C_{\xi\eta} & C_{\xi\xi} \end{bmatrix} \qquad (28)$$

que es función de los elementos de los sistemas de ecuaciones (24, 26) y de los supuestos asociados. Se tiene:

$$C_{\eta\eta} = B^{-1}\Gamma\Phi\Gamma'B'^{-1} \quad B^{-1}\Psi B'^{-1} \qquad (29)$$

$$C_{\eta\xi} = C'_{\xi\eta} = B^{-1}\Gamma\Phi \qquad (30)$$

$$C_{\xi\xi} = \Phi \qquad (31)$$

En el caso general, la matriz de varianzas-covarianzas de las variables observables implicadas en el modelo, es:

$$E(zz') = \Sigma = \left[\begin{array}{c|c} \Lambda_y C_{\eta\eta}\Lambda'_y + \theta_\varepsilon & \Lambda_y C_{\eta\xi}\Lambda'_x \\ (E(yy')) & (E(yx')) \\ \hline \Lambda_x C_{\xi\eta}\Lambda'_y & \Lambda_x C_{\xi\xi}\Lambda'_x + \theta_\delta \\ (E(xy')) & (E(xx')) \end{array}\right] \qquad (32)$$

La importancia de la matriz Σ, no sólo reside en que caracteriza la distribución conjunta de las variables observables de manera que, si se supone a las variables observables con una distribución multivariable normal de media cero, la matriz Σ incluye todos los parámetros que determinan dicha distribución, sino que, además, la ecuación (32) explicita que los elementos de esta matriz, son función de los parámetros del modelo general (24, 26): $\pi = (B, \Gamma, \Phi, \Psi, \Lambda_x, \Lambda_y, \theta_\delta, \theta_\varepsilon))$, es decir, está "estructurada" en términos de otro conjunto de parámetros que deben ser estimados. Esta relación entre matriz de varianzas-covarianzas y parámetros del modelo, permite como se verá en el apart. 5.5 la estimación de éstos últimos.

Los elementos de estas matrices pueden ser, analogamente a lo que ocurre en el modelo AEC, fijos, restringidos o libres.

La ventaja más sobresaliente del sistema general que se presenta, es que las relaciones formuladas en la ecuación (32) no deben ser derivadas por separado para cada modelo particular, sino que, en cada caso, pueden concretarse efectuando las especificaciones apropiadas en este sistema general. Así el modelo más general de medida, AEC, está incluido en (32).

5.4 IDENTIFICACION

Una vez se ha especificado el modelo, y previo incluso a la recogida de datos con los que se calcularán las varianzas y covarianzas muestrales para estimar los parámetros de π, debe examinarse el problema de la identificabilidad.

Los métodos de estimación que se verán en el próximo apartado, proporcionarán estimadores consistentes sólo para aquellos parámetros del modelo que sean identificables. Para exa minar el problema de la identificación, considérense los elementos de la matriz $\sum$ del modelo en la forma general:

$$\sigma_{ij} = f_{ij} (\pi) \quad ; \quad i \leq j \tag{33}$$

donde f_{ij}, como se ha visto en apartados anteriores, es una función contínua no lineal de los parámetros incluidos en π.

Si bien las varianzas-covarianzas, σ_{ij}, pueden siempre derivarse de los valores de los parámetros estructurales de manera única, éstos no pueden obtenerse siempre a partir de las varianzas y covarianzas de las variables observables. El estudio de las condiciones para la unicidad en la determinación de los parámetros se conoce como el problema de la identificación del modelo.

Así pues, los parámetros pueden derivarse de los datos sólo bajo ciertas condiciones, y esto, aún conociendo perfectamente a la población (el problema de inferencia estadística que surge al tener en cuenta que la matriz $\sum$ es en realidad sustituida por la de varianzas-covarianzas muestral S, se tratará en el apartado siguiente). En efecto, existen 1/2 (p+q) (p+q+1) ecuaciones (elementos σ_{ij}) y un número de elementos desconocidos en π, por tanto, una condición necesaria para la identificación de todos los parámetros, es que la diferencia entre ecuaciones e incógnitas (parámetros de valor desconocido), a la que se llama número de grados de libertad ν, debe ser no negativa. Según el valor de ν los modelos pueden clasificarse en:

- Indeterminados ("underidentified"): $\nu<0$, modelo en que los parámetros pueden tomar tantos valores que no están determinados unívocamente. Resulta imposible su contrastación.

- Determinados ("exactly identified"): $\nu=0$, existe una transformación biyectiva entre datos y parámetros. El modelo no puede contrastarse, luego no es susceptible de ser refutado.

- Sobredeterminados ("overidentified"): $\nu>0$, en otras palabras, esto significa que el modelo incluye menor número de parámetros que de varianzas y covarianzas muestrales. Sólo en este caso, un modelo es rechazable en base a los datos, habitualmente varianzas y covarianzas.

Tres categorías que resultan "no simétricas", pues mientras la analogía con la resolución de los sistemas algebraicos de ecuaciones puede ayudar a entender que en el primer caso no existe información suficiente para estimar los parámetros estructurales (que sí podrían ser estimados, en cambio, cuando el número de grados de libertad fuera nulo), no aclara que en los modelos sobredeterminados, a diferencia de lo que ocurre en aquéllos (sistemas algebraicos) no sólo es posible la estimación de los parámetros, sino que las relaciones excedentes permitirán la comprobación del modelo (lo que resulta inviable en cualquier otro caso), pues si el modelo es correcto, estas "ν" ecuaciones deberían ser redundantes, es decir, compatibles con las estimaciones obtenidas a partir de las ecuaciones escogidas para resolver el sistema.

5.5 ESTIMACION.

Resuelto el problema de la identificación, surge el de la estimación de los parámetros desconocidos. En la última sección, se ha ignorado la diferencia entre la matriz de varianzas-covarianzas de la población $\sum$ y la muestral S, en aras a la simplicidad con que quiso enfocarse la identificación del modelo. La introducción de esta distinción es el punto fundamental de este apartado.

Del examen de las ecuaciones (24, 26) del sistema general, se llega a la conclusión de que la población está caracterizada por los parámetros incluidos en π, todos ellos elementos de la matriz $\sum$. En la práctica, estos parámetros serán desconocidos y deberán estimarse a partir de los datos sobre N individuos que se recogen en la matriz de datos X(Nxp). Para esto debe calcularse el vector de medias muestrales $\bar{x}'=(\bar{x}_1,\bar{x}_2,\ldots\ldots,\bar{x}_p)$ y la matriz S de varianzas-covarianzas muestrales (o la de correlaciones R).

Generalmente, el vector de esperanzas matemáticas μ no se restringe tomando simplemente a $\bar{x}$ como estimación de éste. El resto del problema de la estimación consiste, esencialmente, en ajustar la matriz Σ impuesta por el modelo en la ecuación (32) a la de varianzas-covarianzas muestral S que se considera particionada según Σ:

$$S((p+q)x(p+q)) = \left[\begin{array}{c|c} S_{yy}(pxp) & S_{yx}(pxq) \\ \hline S_{xy}(qxp) & S_{xx}(qxq) \end{array}\right] \quad (34)$$

Se consideran a continuación tres de los criterios más habituales para ajustar la matriz Σ de la ecuación (32) a la matriz observada S:

1) Mínimos cuadrados no ponderados (U.L.S.) que minimiza:

$$U = \frac{1}{2}\,\mathrm{tr}(S-\Sigma)^2 \quad (35)$$

2) Mínimos cuadrados generalizados (G.L.S.) que minimiza:

$$G = \frac{1}{2}\,\mathrm{tr}(I-S^{-1}\Sigma)^2 \quad (36)$$

3) Máxima verosimilitud (M.L.) que minimiza:

$$F = \mathrm{tr}(\Sigma^{-1}S) - \ell n|\Sigma^{-1}S|-p \quad (37)$$

Los estimadores obtenidos, cualquiera que sea el criterio utilizado, U,G, ó F, resultan asintóticamente insesgados y consistentes. Además, supuesto el vector de las variables x con distribución multinormal, los criterios G.L.S. y M.L. proporcionan estimadores de varianza mínima. Sin embargo, los dos últimos G y F requieren que la matriz S (ó R) sea definida positiva, mientras que U puede trabajar incluso con matrices no-gramnianas (Véase Jöreskog /22/).

Para ajustar S a Σ, LISREL utiliza el procedimiento de la máxima verosimilitud con las adecuadas modificaciones para tratar modelos con parámetros restringidos y fijos.

La última versión, LISREL V, puede obtener, además, estimadores U.L.S., e igualmente permite tratar con variables no contínuas (véase Jöreskog y Sörbom /74/).

5.6 VERIFICACION.

Se ha sugerido como pueden formularse teorías mediante modelos, que se derivarían del sistema general (24, 26) introduciendo restricciones en los posibles valores de los parámetros del vector π, que definen y especifican un modelo concreto.

Por otro lado, se habrán recogido los datos relevantes en el contexto del diseño apropiado (Véase el esquema de la Fig. 2). En estas condiciones, es posible contrastar el modelo supuesto y los datos recogidos, pues si las restricciones específicadas fueron correctas, el ajuste del modelo a los datos será aceptable, pero incorrectas restricciones (independientemente de la selección de valores que la estimación hiciese) redundarían en el empobrecimiento del ajuste, rechazándose, si procediera, al modelo como plausible representación del proceso causal que genera los datos. Así pues, comprobar la bondad del ajuste puede indicar si el modelo fué correctamente especificado o no, es decir, verificar su validez.

Junto a este test global del modelo, que depende del criterio de estimación elegido (GLS o ML) y que se basa en el estadístico X^2, podrán formularse tests para subconjuntos de hipótesis en el contexto de una teoría dada. Estos tests se basarán en la comparación del ajuste que proporcionarán pares de modelos, (véase tabla 3), por cuanto la diferencia de ajuste entre ambos modelos se imputará (explicará) al conjunto de hipótesis que se deseen contrastar.

Por último, los criterios de estimación de la máxima verosimilitud y de los mínimos cuadrados generalizados permitirán comprobar la significación individual de cada parámetro, ya que al obtener la inversa de la matriz de información pueden calcularse las varianzas de los parámetros estimados (Jöreskog /27/. Ver tabla 4).

LISREL V proporciona medidas adicionales para comprobar la bondad del ajuste, tales como el coeficiente de correlación múltiple para las variables de interés y los coeficientes de determinación tanto para las ecuaciones de medida como para las estructurales.

Para mayor información sobre estas etapas del modelo véase Jöreskog /69/, Saris y Stronkhorts /63/, Jöreskog y Sörbom /73/, /74/.

5.7 EJEMPLOS ILUSTRATIVOS.

Saris y Stronkhorts /63/, quieren establecer el proceso causal que determina la selección de la escuela secundaria en Holanda. El diagrama "path" que refleja dicho proceso es:

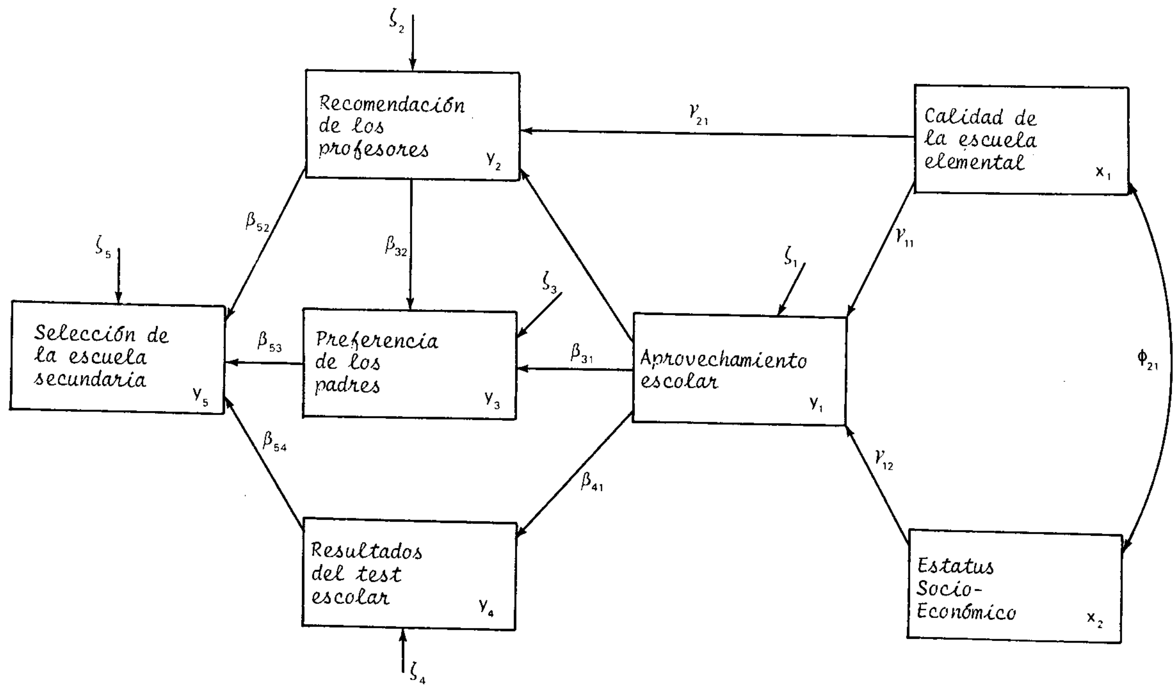

Fig. 3 Modelo propuesto inicialmente para el proceso de selección de la escuela secundaria /63/.

Como se desprende de la tabla 3, este modelo teórico de partida es rechazado. El proceso de optimización del modelo se basa en la contrastación jerárquica y secuencial a la que se refería el apartado 5.6 y se lleva a cabo añadiendo aquellos parámetros que influyen en los residuos (diferencia entre covarianza observada y reproducida a partir del modelo; ver tabla 2) y cuyas primeras derivadas (que indican la dirección del vector gradiente -- utilizable para minimizar la función (37)) sugieran su inclusión en el modelo, salvo que contradigan la teoría subyacente al mismo. - Las tablas 1 y 2 expresan el valor de dichos índices en la introducción del primer parámetro, γ_{41}, al modelo original. El proceso global de contrastación que conduce a la adopción del modelo final se resume en la tabla 3.

Una vez se ha obtenido el modelo "que ajusta" (10º) puede aplicarse el test de significación de los parámetros individuales del apartado 5.6. Las conclusiones a las que llega se han recogido en la tabla 4.

Asi, el modelo que en definitiva representa el proceso de selección de la Escuela secundaria en Holanda es, según Saris y Stronkhorts /63/:

$$
\begin{aligned}
y_1^* &= .285x_1^* + .224x_2^* + \zeta_1^* \\
y_2^* &= .784y_1^* + .123x_2^* + \zeta_2^* \\
y_3^* &= .168y_1^* + .711y_2^* + .077x_1^* + .050x_2^* + \zeta_3^* \\
y_4^* &= .431y_1^* + .170y_2^* + .224y_3^* + .225x_1^* + \zeta_4^* \\
y_5^* &= .035y_2^* + .933y_3^* + .035y_4^* + \zeta_5^*
\end{aligned} \tag{38}
$$

en donde el asterisco (*) significa que las variables han sido estandarizadas.

El diagrama causal equivalente al sistema de ecuaciones (38) es:

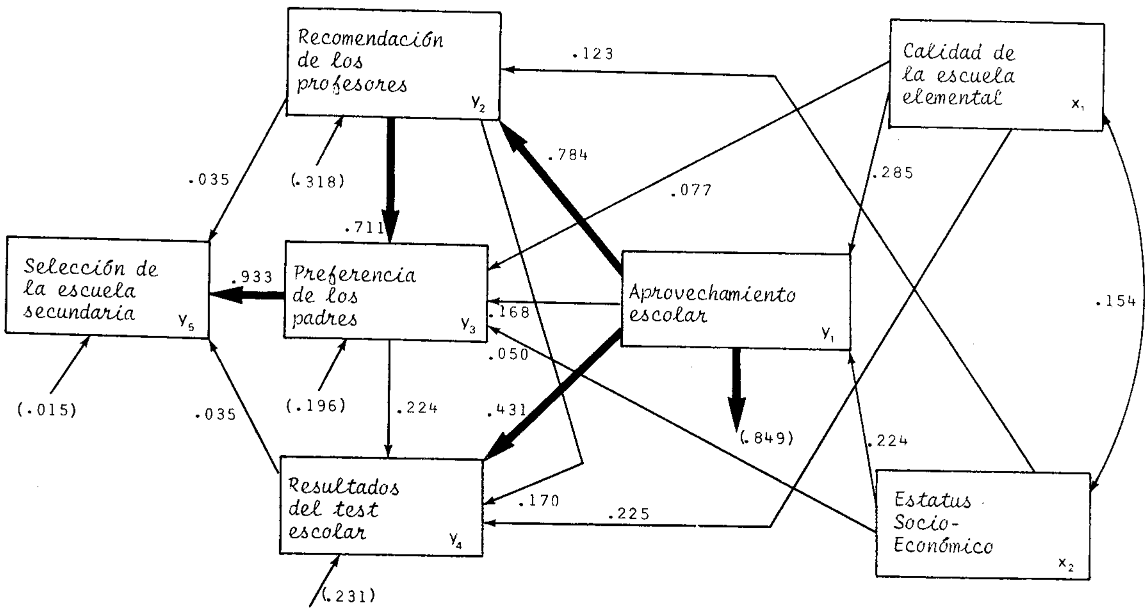

Fig. 4 Diagrama causal del proceso de Selección de la Escuela secundaria en Holanda.

Todo ello implica, en el sentido del "Path - Analysis" /28/, la descomposición en los --- efectos causales indicada en la Tabla 5.

Otras aplicaciones de LISREL pueden verse en /23/ y /75/.

TABLA 1: VALOR QUE TOMAN LAS PRIMERAS DERIVADAS EN EL MODELO INICIAL DE LA FIG. 3

DERIVADAS DE PRIMER ORDEN

BETA

	1	2	3	4	5
1	-.000	.030	.045	.094	.025
2	-.000	.000	-.018	-.333	-.032
3	.000	.000	.000	-.297	-.012
4	.000	-.336	-.432	.000	-.413
5	.082	-.000	-.000	.000	-.000

GAMMA

	1	2
1	-.000	-.000
2	.000	-.344
3	-.340	-.000
4	-.662	-.232
5	.444	.054

PHI

	1	2
1	-.000	
2	-.000	-.000

PSI

	1	2	3	4	5
1	-.000				
2	.091	.000			
3	.114	.004	.000		
4	.284	-1.004	-.896	.000	
5	-.067	-.206	-.093	-.202	-.000

TABLA 2: MATRIZ DE LOS RESIDUOS PARA EL MODELO DE LA FIG. 3

RESIDUOS : S - SIGMA

	Y1	Y2	Y3	Y4	Y5	X1
Y1	0.000					
Y2	0.000	0.000				
Y3	-0.000	0.006	0.009			
Y4	0.000	0.111	0.143	0.000		
Y5	-0.001	0.010	0.014	0.138	0.018	
X1	0.000	0.000	0.069	0.219	0.065	0.000
X2	0.000	0.114	0.081	0.077	0.082	0.000

RESIDUOS : S - SIGMA

	X2
X2	0.000

TABLA 3: CONSECUCION DEL MODELO FINAL A TRAVES DEL PROCESO DE INCLUSION SECUENCIAL DE LOS PARAMETROS

Modelo	Descripción del modelo	Valor del estadístico	g.l. (ν)	Nivel de significación	Diferencia en la bondad del ajuste	Diferencia en ν	Aceptación de la corrección	Proporción en la mejora
(1º)	Independencia mutua	4471.07	23	.0000	-	-	-	-
(2º)	Modelo de la Fig. 3	225.22	9	.0000	3255.85	4	si	.950
(3º)	Modelo (2º)+γ_{41}	134.11	8	.0000	91.11	1	si	.970
(4º)	Modelo (3º)+γ_{51}	132.06	7	.0000	2.05	1	no	.970
(5º)	Modelo (3º)+γ_{22}	111.63	7	.0000	21.48	1	si	.975
(6º)	Modelo (5º)+γ_{31}	97.92	6	.0000	13.71	1	si	.978
(7º)	Modelo (6º)+γ_{42}	90.57	5	.0000	7.35	1	si	.980
(8º)	Modelo (7º)+ +γ_{51}+γ_{52}	88.50	3	.0000	2.07	2	no	.980
(9º)	Modelo (7º)+β_{43}	13.19	4	.0104	77.38	1	si	.999
(10º)	Modelo (9º)+β_{42}	2.36	3	.5005	10.83	1	si	.999

TABLA 4: TEST DE HIPOTESIS PARA LOS COEFICIENTES ESTIMADOS DEL MODELO AJUSTADO (10º)

Parámetro	Estimación puntual	"Standard error"	Intérvalo de confianza: de	a
γ_{11}	.285	.041	.205	.365
γ_{12}	.224	.041	.144	.304
β_{21}	.786	.027	.733	.839
γ_{21}	-.006	.026	-.057	.045
γ_{22}	.124	.026	.073	.175
β_{31}	.168	.034	.101	.235
β_{32}	.711	.035	.642	.780
γ_{31}	.077	.021	.036	.118
γ_{32}	.050	.021	.009	.091
β_{41}	.432	.038	.394	.470
β_{42}	.167	.051	.067	.267
β_{43}	.221	.048	.127	.315
γ_{41}	.224	.023	.179	.269
γ_{42}	.014	.023	-.031	.059
β_{52}	.035	.012	.011	.059
β_{53}	.933	.012	.909	.957
β_{54}	.035	.009	.o17	.053

TABLA 5: DESCOMPOSICION EN EFECTOS CAUSALES

Variable causal	Efecto directo en y_5	Efecto indirecto en y_5	Efecto total en y_5
Resultado del test (y_4)	.035	.000	.035
Preferencia de los padres (y_3)	.933	.011	.944
Recomendación de los profesores (y_2)	.035	.674	.709
Aprovechamiento escolar (y_1)	.000	.729	.729
Calidad de la escuela primaria (x_1)	.000	.288	.288
Estatus Socio-económico (x_2)	.000	.298	.298

6. CONCLUSIONES Y OTROS ASPECTOS.

La metodología exploratoria del AF ha ido evolucionando hacia un enfoque confirmatorio. Las diferentes generalizaciones de este enfoque que conducen al AEC y al planteamiento LISREL, que explican una gran variedad de modelos multivariables de aplicación corriente en Psicología y otras ciencias sociales y biológicas. La principal aportación del AEC y LISREL en la teoría de los modelos causales, consiste en conjugar las relaciones medidas por las covarianzas en las relaciones entre causas y efectos, con la posibilidad de investigar el modelo más adecuado mediante las etapas de identificación, estimación y verificación, dentro de la objetividad que es inherente a la metodología de la inferencia estadística. En líneas generales, se trata de analizar hasta que punto la matriz de covarianzas observada se ajusta a la estructura que se deduce de un modelo planteado. Quizás, la principal aportación del planteamiento LISREL es que permite comprender y justificar la causalidad en situaciones no reproducibles experimentalmente, siendo innumerables las aplicaciones que tiene en el campo de las ciencias sociales y de la conducta.

La introducción de factores o causas latentes, presenta algunos problemas, entre los que destacamos:

1) Al ser un factor una variable no observable puede llegar a tener un significado esencialmente especulativo.

2) Diferentes medidas en los factores pueden verificar coherentemente un mismo modelo factorial.

3) La relación entre factores y variables externas (es decir, variables no incluidas en la batería de tests en el caso del AF) es indeterminada.

La correlación entre variables externas y factores ha sido estudiada por Cuadras /76/, Cuadras y Sánchez /77/. Steiger /78/ obtiene una acotación de estas correlaciones.

Finalmente, cabe tener en consideración los aspectos siguientes, no plenamente resueltos:

a) El modelo que relaciona variables observables con factores se justifica a partir de la estructura de la matriz de covarianzas. Sin embargo, esta estructura no implica necesariamente un determinado modelo de AF, AEC ó LISREL (por ejemplo: (11) no implica (10), aunque (10) implica (11)),

b) El efecto de las desviaciones de la normalidad sobre la estimación de parámetros y tests de ajuste,

c) El caso de variables ordinales o cualitativas,

d) Planteamientos no paramétricos del problema,

e) Modelos no lineales coherentes e interpretables,

f) La comparación de modelos,

Finalmente, es de destacar la reciente aportación de Satorra /79/ en el estudio de la función de potencia en tests de hipótesis sobre LISREL, analizando la influencia del tamaño muestral, de las desviaciones de la hipótesis, de los errores de especificación, etc.

7. AGRADECIMIENTO.

Agradecemos al profesor W.Saris la introducción al LISREL, así como la correspondencia mantenida durante estos tres últimos años, que nos ha sido de inestimable valor.

Agradecemos también las valiosas sugerencias de A. Satorra y C. Molinas que han permitido mejorar algunos aspectos de este trabajo.

8. BIBLIOGRAFIA.

/1/ JOLICOEUR, P. y MOSIMANN, J.E. "Size and shape variation in the painted turtle: A principal component analysis", Growth, 24 (4)(1960), 339-354.

/2/ BURNABY, T.P., "Growth-invariant discriminant functions and generalized distances", Biometrics, 22 (1)(1966), 96-110.

/3/ PEARSON, K. "On lines and planes of closest fit to a system of points in space", Phil.Mag., 2, 6th series (1966), 151-180

/4/ HOTTELLING, H. "Analysis of a complex of statistical variables into principal components", J. Educ.Psychol, 24 (1933), -- 417-441, 498-520.

/5/ HOTTELLING, H. "The most predictable criterion", J. Educ. Psychol., 26 (1935), - 139-141.

/6/ LEBART, L., MORINEAU, A, TABARD, N., --- "Techniques de la description statistique", Dunod. (1977)

/7/ ESTRADA, M., BLASCO, D. "Two phases of - the Phytoplankton Community in the Baja California Upwelling" Limnol. Oceanogr., 24(6), (1979), 1065-80.

/8/ ZUNICA, L.R. "Características socioeconómicas y comportamiento electoral, el caso valenciano", QUESTIIO 5, 4 (1981) 189-202.

/9/ CUADRAS, C.M., OLLER, J.M., "Mètodes de representació de dades i altres tècniques estadístiques aplicades a les ciencies sociàls", Les eleccions de 1977 a Catalunya. Estudis Electorals". Pub. Fund. J. Bofill (1981), 421-448

/10/ BERTIER, P., BOUROCHE, J.M., "Analyse -- des donneés multidimensionnelles" Presses Universitaries de France, Paris ---- (1975).

/11/ BATISTA, J.M., ESTIVILL, X., "Aplicació del analisi de components principals a nivell de dades Municipals: Definició - de zones homogenees i estudi de la jerarquia urbana a Catalunya" (en prensa).

/12/ SPERRMANN, CH., "General Intelligence, objectively determined and measured" -- Amer. J. Psychol, 15 (1904), 201-293.

/13/ KAISER, H.F., "The varimax criterion for analytic rotation in factor analysis" - Psychometrika, 23 (1958), 187-200.

/14/ THURSTONE,L.L., "Multiple factor analysis", Psychol. Revv., 38 (1931),406-427.

/15/ HARMANN, H.H., "Modern factor analysis" Chicago, University of Chicago Press. - (1976).

/16/ CUADRAS, C.M., "Métodos de analisis multivariante", EUNIBAR (1981).

/17/ KAISER, H.F., CAFFREY, J., "Alpha factor analysis", Psychometrika, 30 (1965),1-14.

/18/ SWAMINATHAN, H., ALGINA, J.,"Scale freeness in factor analysis", Psichometrika, 43 (1978) 581-583.

/19/ AMSTRONG, J.S., "Derivation of Theory by means of factor analysis or Tom Swift and his electric factor analysis machine". - The American Statistician, (1967), 17-21.

/20/ BOCK, R.D., "Components of variance analysis as a structural and discriminal analysis for psychological tests", British - Journal of Statist. Psychol., 8 (1960), - 151-163.

/21/ BOCK, R.D., BARGMANN, R.E.- "Analysis of covariance Structures", Psychometrika,31, (1966)507-534.

/22/ JÖRESKOG, K.G.. "A general method for -- analysis of covariance structures", Biometrika, 57 (1970), 239-251.

/23/ BATISTA, J.M., "Modelo para el análisis de ecuaciones de estructura lineal, LISREL. Relación con el modelo para el análisis de la estructura de la covarianza-aplicación a un programa preventivo frente al consumo de drogas en adolescentes". Tesis Doctoral. Universidad Politécnica de Barcelona (1982).

/24/ JÖRESKOG, K.G., "Factor analysis by least squares and maximum likelihood methods". En: Statistical Methods for Digital Computers. (K.Enslein, A. Ralston y H.S. Wilf, Eds.), J. Wiley, New-York, (1977).

/25/ JÖRESKOG, K.G., GRUVAEUS,G.T., y VAN --- THILLO, M. " A General Computer Program for Analysis of Covariance Structures", Research Bull. 70-15. Princenton, N.J.; Educational Teskins Services, (1970).

/26/ GOLDBERGER, A.S., DUNCAN, O.P., "Structural Equation Models in Social Scien--ces", Seminar Press, (1973).

/27/ PENROSE, L.S., " Some notes of discrimination" Ann. Eugen, 13 (1947), 228-237.

/28/ JÖRESKOG, K.G., "A general approach to confirmatory factor analysis", Psychometrika, 34 (1969), 183-202.

/29/ DUNCAN, O.D.,"Introduction to structural equation models" New York Academic ----- (1975).

/30/ HOWE, W.G., "Somo contributions to factor analysis" Report ORNL-1919, Oak Ridge National Laboratory, Oak Ridge, Tenesee (1975).

/31/ ANDERSON, T.W., RUBIN, H., "Statistical inference in factor analysis", Proc. of the third Berkeley Symposium on Mathem. Stat. and Probability, 5, (1956) 111-50.

/32/ LAWLEY, D.N., JÖRESKOG, K.G., "New --- methods in maximum likelyhood factor -- analysis", British Journal of Math. and Statist. Psychol., 21, (1968), 85-96.

/33/ JÖRESKOG, K.G., "Structural analysis of covariance and correlation matrices", - Psychometrika 43, (1978), 443-477.

/34/ MULAIK, S.A., "The Foundations of Factor Analysis", New York : Mc.Graw-Hill ---- (1972).

/35/ LAWLEY, D.N. y MAXWELL, A.E., "Factor -- Analysis as Statistical Method", London Butterworth (1971).

/36/ JÖRESKOG, K.G. y SÖRBOM, D., "Confirmatory factor analysis with model modifica--tion" (COFFAM) International Educational Services, Chicago Illinois 60615.

/37/ LORD, F. y NOVICK, M., "Statistical theories of mental test scores", Reading --- Addison Nesley, (1969).

/38/ JÖRESKOG, K.G., "Statistical analysis of sets of Gongeneric tests", Psychometrika 36 (1971) 109-133.

/39/ KRISTOF, W., "On the theory of a set of tests which differ only in length" Psychometrika, 36, (1971) 207-225.

/40/ ALWIN, D.F., JACKSON, D.J., "Measurement models for response errors in surveys: Issues and application". Sociological -- Methodology, (1980).

/41/ JANSEN, A.M.M. "Comparative calibration and congeneric measurements", Biometrics 36 (1980), 117-135.

/42/ SARIS, W.E., "Different questions, different variables?" Report of the Department of Methodology Free University Amsterdam (1979).

/43/ CRONBACH, L.J., GLESER, G., RAJARATNAM, N., "Dependatibility of behavioral measurement. Theory of generability for -- scores and profiles", Wiley, New York, (1972).

/44/ WILEY, D.E., SCHMIDT, W.H. y BRAMBLE,W.J. 'Studies of a class of covariance structure models", J.A.S.A., 68, (1973) 317-323.

/45/ MULAIK, S.A., "Confirmatory factor analysis", En Amick y Walberg (Eds.):Introductory Multivariate Analysis Berkeley: --- McCutchan, (1975).

/46/ JÖRESKOG, K.G., "Analysis of covariance structures" En Krishnaiah (Ed.): Multivariate Analysis III. New York, Academic Press, (1973).

/47/ MALLENBERG, KELDERMAN, STIJLEN, ZONDAG "Linear models for the analysis and construction of instruments in a facet de--sign" Colleg.Bull., 86, 4 (1979), 766--776.

/48/ DEKKIN y RAADSHEER, "Konstruktie vaneen instrument voorhetmeten van sociale --- angst big kinderen" (sin publicar), Universidad de Amsterdam, Laboratorio de - Psicologia (1977).

/49/ SARIS, W.E., "Linear structural relationships. Measurement models" Vol.II, Free University. Amsterdam (1982).

/50/ CAMPBELL, D.T., FISKE, D.W. "Convergent and discriminant validation by the multitrait-multimethod matrix", Psychol -- Bulletin 56 (1959) 81-105.

/51/ JÖRESKOG, K.G., "Analyzing-psychological data by structural analysis of covariance matrices", En Contemporany develop--- ments in math. psychol. Atkinson y otros (Ed.). San Francisco . Freeman (1974), - 1-56.

/52/ WERTS, C.E., JÖRESKOG, K.G. y LINN, R.L. "A multitrait multimethod model for studying growth". Educational and Psychological Measurement, 32 (1972), 655-678.

/53/ ALWIN, D.F., "Approaches to the interpretation of relationships in the multi--- trait multimethod matrix". En Costner -- (Ed) Sociological methodology 1973-1974. San Francisco, Jossey-Bass.

/54/ SARIS, W.E., NEIJENS,P, y DOORN, L.V., - "Scaling social sciences variables by -- multimodality matching", M.D.N., 2 (1980) 3-22.

/55/ GUTTMAN, L., "A new approach to factor - analysis: The Radex"., En: Lazarsfeld -- (Ed), Mathematical thinking in the so--- cial sciences New York. Columbia University Press, (1954).

/56/ ANDERSON, T.W., "Some stochastic process models for intelligence test scores".En: Arrow, Karlin y Suppes (Ed)., Mathematical methods in the social sciences, Stanford University Press, (1960).

/57/ WRIGHT, J., "The method of path coeffi-- cients" Annals of Math. Statistics, 5 -- (1934), 161-215.

/58/ WRIGHT, J., "Path coefficients and path regressions: Alternative or complementary concepts?". Biometrics, 16 (1960), -- 189-202.

/59/ WERTS, C.E. y LINN, R.L., "Path analy-- sis: Psychological examples", Psychol. - Bulletin, 74 (1970) 194-212.

/60/ HAUSER, R.M. y GOLDBERGER, A.S., "The -- treatment of unobservable variables in - path analysis". Costner (Ed). Sociological Methodology. San Francisco: Jossey Bass, (1971), 81-117.

/61/ JÖRESKOG, K.C. y GOLDBERGER, A.S.,"Estimation of a model with multiple indicators and multiple causes of a single latent variable", J.A.S.A., 10, (1975), - 631-639.

/62/ ARNAU, J., "Psicologia experimental (un enfoque metodológico)", Trillas, Mexico, (1979).

/63/ SARIS, S. y STRONKHORTS, H., "Introduction to causal models in non-experimental research" Free University of Amsterdam y Central Bureau of Statistics. --- Voorburg, (1981).

/64/ BENTLER, P.M., "Multivariate analysis - with latent variables", Ann. Rev. Psy-- chol., 31 (1980), 419-456.

/65/ SIMON, H.A., "On the definition of the causal relation", Journal of Philosophy, 49, (1952), 517-528.

/66/ BLALOCK, H.M., "Causal models in the social sciences". Chicago. Adine-Atherton, (1971).

/67/ BOUDON, R., "A method of linear causal analysis: Dependence analysis", AM. Sociol. Rev. 30, (1965), 365-373.

/68/ DUNCAN, O.D., "Path analysis: Sociological examples", A.J.S., 72, (1966), 1-16.

/69/ JÖRESKOG, K.G., "A general method for - estimating: A linear structural equation system", Goldberger y Duncan (1973), 85-112.

/70/ KEESLING, W., "Maximum likelihood approaches to causal flow analysis", PhD.Thesis Univ. Chicago. Chicago, Illinois, (1972).

/71/ WILEY, D.E., "The identification problem for structural equation models with un-- measured variables", Goldberger y Duncan (1973), 69-83.

/72/ JÖRESKOG, K.G. y VAN THILLO , M., "LISREL - A general computer program for estimating a linear structural equation - system involving multiple indicators of unmeasured variables". Department of -- Statistics, University of Upsala,(1974).

/73/ JÖRESKOG, K.G. y SÖRBOM, D., "LISREL IV - A general computer program for esti-- mation of a linear structural equation system by maximum likelihood methods", Chicago: National Educational Services (1978).

/74/ JÖRESKOG, K.G.,y SÖRBOM, D., "LISREL V, Analysis of linear structural relation- ships by maximum likelihood and least squares methods", Department of Statis- tics. University of Upsala, (1981).

/75/ MARUYAMA, G.. McGARVEY,B, "Evaluating causal models: An application of maxi- mum-likelihood analysis of structural equations", Psycho. Bulletin, 87(3), - (1980), 502-512.

/76/ CUADRAS, C.M., "Sobre la reducción de la dimension en analisis estadístico - multivariante", Trab. Est. I. Oper. -- 28(1), (1977), 63-76.

/77/ CUADRAS, C.M., SANCHEZ, M., "The con-- cept of a common factor in factorial - statistical analysis and its relation with the external variables", Proc. -- First World Conf. Mathem. Serv. Man. - Vol. I, (1979), 531-540.

/78/ STEIGER, J.H., "The relationship bet-- ween external variables and common fac tors", Psychometrika, 44(1), (1979), - 93-98.

/79/ SATORRA, A., "Potència del contrast de la raó de versemblança en models d'equa cions estructurals", Tesi doctoral (inè dita, F. de Matemàtiques, Univ. de Bar- celona, (1983).