

VI Congreso Chileno de Antropología. Colegio de Antropólogos de Chile A. G, Valdivia, 2007.

El Análisis Factorial Múltiple: Un Paso más en la Superación de la Dicotomía Cualitativo- Cuantitativo.

Juan Moncada Herrera.

Cita:

Juan Moncada Herrera (2007). *El Análisis Factorial Múltiple: Un Paso más en la Superación de la Dicotomía Cualitativo-Cuantitativo. VI Congreso Chileno de Antropología. Colegio de Antropólogos de Chile A. G, Valdivia.*

Dirección estable: <https://www.aacademica.org/vi.congreso.chileno.de.antropologia/67>

ARK: <https://n2t.net/ark:/13683/eCzH/Dvz>

Acta Académica es un proyecto académico sin fines de lucro enmarcado en la iniciativa de acceso abierto. Acta Académica fue creado para facilitar a investigadores de todo el mundo el compartir su producción académica. Para crear un perfil gratuitamente o acceder a otros trabajos visite: <https://www.aacademica.org>.

El Análisis Factorial Múltiple: Un Paso más en la Superación de la Dicotomía Cualitativo-Cuantitativo¹

Multiple Factorial Analysis: A Step Further in the Overcoming of the Dichotomy Qualitative-Quantitative

Juan Moncada Herrera*

Resumen

En Investigación Social, uno de los problemas de las técnicas estadísticas clásicas, tales como el Análisis Factorial, el Análisis en Componentes Principales, el Análisis de Correspondencias, etc., es que no toman en cuenta la estructura de grupos que muchas veces tiene la información. En este trabajo se presenta y discuten algunas posibilidades del Análisis Factorial Múltiple (AFM) en esta dirección AFM, enfatizando que la técnica puede ser un importante medio de vinculación de lo cualitativo y lo cuantitativo.

Palabras Claves: Tablas Múltiples, Análisis Factorial Múltiple, indicadores de pobreza.

Abstract

In Social Research, one of the problems of classical statistics techniques, such as the Factorial analysis, Principal Components Analysis, Correspondence Analysis, etc., is that it does not consider the structure of groups that often have the information. In this work, we present and discuss some possibilities of Multiple Factorial Analysis (MFA) in this direction MFA, emphasizing that the technique can be an important means of link between the qualitative and quantitative.

Keywords: Multiple tables, Multiple factor analysis, poverty indicators.

I. Introducción

En la medida que el ser humano progresa, avanza y evoluciona en los múltiples ámbitos y facetas de su vida, tanto individual como colectiva, en materia de información y conocimiento se dan dos fenómenos que merecen ser atendidos sin que sobre uno prime cierta importancia. Se trata, por un lado, de una suerte de 'acu-

mulación' de información que requiere ser organizada, tratada, estudiada y analizada, a menudo con cierta urgencia, a los efectos de volver o retornar a las fuentes en forma de conocimiento o tecnología que dé nuevos impulsos a ese crecimiento humano y social natural en todos nosotros. Por otro lado, esa misma información, para transformarse en nuevos insumos tecnológicos o en nuevo conocimiento, requiere de técnicas, métodos y herramientas cada vez más sofisticadas, y que lleven a cabo su cometido sin tener que entrar en importantes pérdidas de información, de confianza, precisión y certeza. Es en este contexto en el que la Estadística Multivariante juega hoy día un papel quizá no dimensionado hace tan sólo algunas décadas. Esta demanda se ve complejizada aún más por la enorme diversidad de temáticas que preocupan en nuestros tiempos en términos de tratamiento y análisis de información. Hace unos cincuenta años tampoco se sospechaba de los enormes avances de la informática, que como bien sabemos, ha venido a proveer a la Estadística de uno de los recursos más importantes para lograr sus diversos cometidos.

El presente trabajo recoge y resume los aspectos más relevantes de una de las técnicas estadísticas pensadas y diseñadas para el tratamiento de grandes volúmenes de información: el Análisis Factorial Múltiple, AFM.

En una primera parte, el trabajo sintetiza los rasgos y aspectos más relevantes del estudio de información multivariante. Seguidamente se presentan antecedentes generales del AFM, sin entrar en detalles de tipo formal, pues sólo se trata de un primer acercamiento, muy intuitivo, a esta técnica estadística de tratamiento

¹ Universidad Católica de Temuco. Facultad de Ingeniería, Dpto. de Cs. Matemáticas y Físicas; Casilla 15-D, Temuco, Chile. E-mail: jmoncada@uctemuco.cl.

de información de tablas múltiples. Finalmente se concluye este trabajo con una aplicación del AFM a un estudio particular sobre pobreza. La idea de esta última parte es ilustrar el uso y utilidad de la técnica y no revelar aspectos desconocidos del tema de estudio. Es más, se espera que el conocimiento generalizado del tema ayude a valorar la utilidad del AFM.

II. La información multivariante y su análisis

En el Análisis Multivariante, como es bien sabido, se trata de estudiar la información contenida en tablas del tipo Individuos x Variables. La mayor parte de los métodos estadísticos existentes se diferencian en función de la naturaleza de las variables y de los objetivos de estudio. Así, por ejemplo, si las variables son de naturaleza cuantitativa, los métodos de Análisis en Componentes Principales, Análisis Factorial, etc., son herramientas útiles en cuanto al estudio de la tabla de datos. Por el contrario, si las variables son cualitativas métodos como el Análisis Factorial de Correspondencias, Análisis Discriminante y muchos otros se podrían seleccionar para aplicar a la tabla de datos.

1. Acerca de la estructura de los datos

Se hace cada vez más importante en el campo de la Estadística Multivariante otra clasificación u óptica, aparte de la naturaleza, con la que tratar la información. Se trata de distinguir o diferenciar grupos de variables, si acaso existen, y de aplicar, sobre esa base de diferenciación informativa, los métodos más adecuados y pertinentes. En este sentido, las variables pueden presentarse, básicamente, en tres formas:

- **Estructura singular.** Hablamos de este tipo de estructura cuando no es posible diferenciar agrupamientos naturales entre las distintas variables. Es decir, todas las variables son homogéneas en relación al criterio que motiva el estudio y los respectivos análisis. Sobre este tipo de datos se suelen aplicar los métodos estadísticos clásicos (AF, AFC, ACP, etc.)
- **Estructura de partición.** Hablamos de este tipo de estructura cuando los elementos de la tabla (variables o individuos) están organizados y distribuidos en grupos, conocida también como **tablas múltiples**. Este tipo de estructura en los datos forma parte de lo que KIERS (1988, 1991) denomina *Datos de conjuntos múltiples*. Se ca-

racteriza porque uno de los tres modos correspondientes está compuesto por conjuntos o clases.

- **Estructura jerárquica.** En muchas situaciones prácticas las variables están estructuradas de acuerdo a una jerarquía, lo cual significa la presencia de conjuntos de variables y *subvariables* dentro de un mismo estudio. Es el caso, por ejemplo, de estudios en los cuales interviene grupos de paneles de evaluadores; o lo que ocurre en encuestas de opinión, en las que la opinión, a su vez, se puede clasificar según alguna otra variable. Una técnica que toma en cuenta esta estructura jerárquica en las variables es el Análisis Factorial Múltiple Jerárquico (LE DIEN & PAGÈS, 2003).

De entre todas estas estructuras, quizá la de partición sea la forma más habitual de encontrar y disponer información. E incluso tal estructura puede encontrarse en las unidades observación (individuos estadísticos), aunque los métodos disponibles para el análisis de tales datos son un poco más complejos y requieren de más desarrollo. En este trabajo la estructura de partición se entenderá referida a variables.

2. Estudios de tablas con estructura de partición

La estructura de partición inducida en las variables, o más sencillamente el estudio de tablas múltiples, conlleva un interés práctico de grandes posibilidades de exploración, ya que posibilita el estudio de la variabilidad en el contexto de los métodos clásicos de análisis de la varianza. En efecto, según la fórmula de Huygens, la inercia total de una nube de puntos se puede descomponer en *inercia inter-clase* e *inercia intra-clase*, en la forma siguiente:

$$\text{Inercia Total} = \text{Inercia inter} + \text{Inercia intra}$$

Esta descomposición da lugar a dos análisis: el *análisis inter-clase* y el *análisis intra-clase*.

El *análisis inter-clase*, que describe las posiciones relativas de las clases sin referencia a los individuos, no es más que un análisis de la tabla agregada de orden (p, q) . En este sentido puede relacionarse con el Análisis de Correspondencias, con el Análisis Factorial Discriminante, con el Análisis Discriminante Baricéntrico, etc.

Por su parte, el *análisis intra-clase*, que se orienta a la descripción o puesta de relieve de las principales características al interior de cada clase, sin consideración del efecto de la estructura de la clase, permite estudiar las diferencias de comportamiento al interior de ella mediante el estudio de la dispersión de los individuos al interior de sus clases de pertenencia.

Los métodos de análisis referidos a tablas con este tipo de estructura se enmarcan –y así deben hacerlo, dada la naturaleza multidimensional de la información en ellas contenida– en la familia de los métodos factoriales.

3. Métodos de análisis de datos de conjuntos múltiples

Presentaremos brevemente los principios generales de los métodos que abordan el tratamiento de Tablas Múltiples mediante la utilización de la DVS de Dos Vías *yuxtaponiendo* las K matrices de datos iniciales, en el contexto de las propuestas de la escuela francesa. En esta propuesta de análisis, las K matrices son ponderadas antes de ser yuxtapuestas. Los factores de ponderación pueden provenir de un análisis factorial clásico aplicado sobre cada una de las matrices iniciales, o bien del análisis de las matrices de productos cruzados obtenidas de cada una de las matrices originales. En este sentido, las matrices yuxtapuestas son matrices de Dos Vías, y su análisis utiliza, en consecuencia, la DVS de Dos Vías.

Los métodos que utilizan esta descomposición, consisten o se estructuran, básicamente, en las siguientes etapas:

- **Primera etapa: Estudio de la INTER-ESTRUCTURA.** Consiste en una comparación global de la estructura de las diferentes tablas. Se define una distancia entre objetos o configuraciones (elementos representativos) que representen a las tablas iniciales, y luego se encuentra una representación euclídea de las tablas, sobre la cual la proximidad entre dos puntos se interpreta en términos de similitud o parecido entre dichas tablas, respecto de la métrica utilizada.
- **Segunda etapa: Búsqueda (o construcción) de un Compromiso.** Consiste en sintetizar to-

das las K tablas en una sola, denominada 'compromiso', que se constituirá en una síntesis o resumen global del conjunto de tablas o matrices.

- **Tercera etapa: Estudio de la INTRA-ESTRUCTURA.** El objetivo de esta etapa es representar, en el 'compromiso' obtenido en la etapa anterior, las 'posiciones compromiso' de cada uno de los elementos (individuos y/o variables) que conforman las distintas tablas. La posición relativa de cada elemento corresponde a la posición media de los mismos.

Entre los métodos que utilizan este esquema de análisis destacan el STATIS, STATIS DUAL y el Análisis Factorial Múltiple (AFM), todos en el contexto de la Escuela Francesa de Análisis de Datos.

III. El Análisis Factorial Múltiple

Nacido al alero de la escuela francesa de análisis de datos, el Análisis Factorial Múltiple (ESCOFIER & PAGÉS, 1992) permite un análisis comparativo de un grupo de tablas referidas a los mismos individuos. De esta manera, su objetivo es estudiar la estructura común de las distintas tablas de datos que constituyen el estudio, proporcionando información sobre los individuos homogéneos, así como también acerca de los elementos heterogéneos. En otras palabras, y de forma similar a cualquier técnica factorial, el AFM persigue poner de relieve los principales factores que dan cuenta de la variabilidad de los individuos, descritos estos, de manera equilibrada, por los diferentes grupos de variables que definen la estructura de partición de la información. El equilibrio de la influencia de los distintos grupos en el análisis se logra mediante una ponderación de las diversas tablas, y de esa manera se posibilita un estudio pertinente de:

- Relaciones entre los distintos grupos de variables, además de medir el grado de semejanza entre tales grupos.
- Relaciones entre las variables que conforman un determinado grupo y las del resto de los grupos.
- Semejanzas entre los individuos, observados estos desde las ópticas de los diferentes grupos de variables.

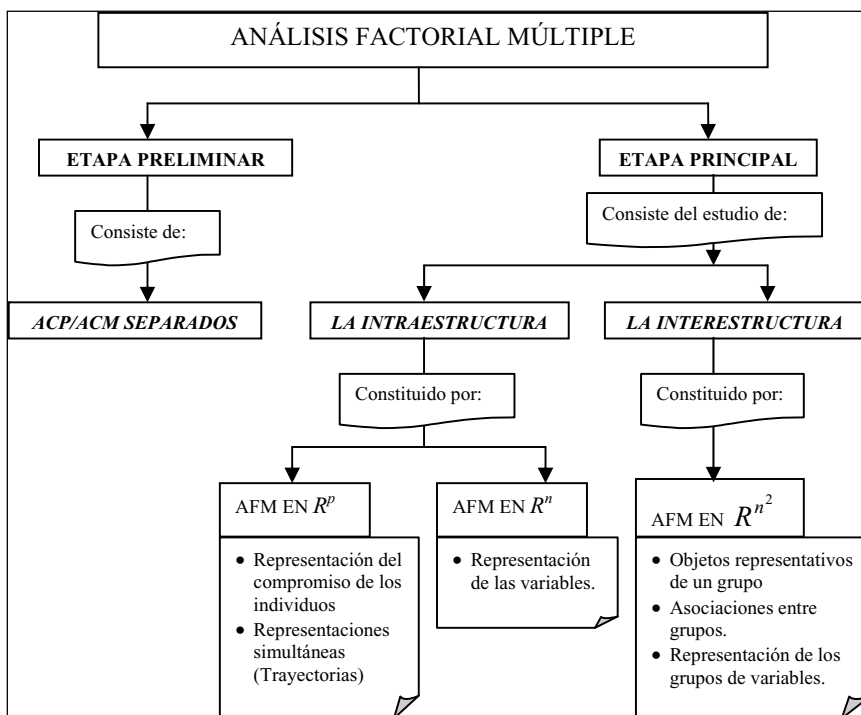


Figura 1: Esquema general de funcionamiento del AFM

1. Funcionamiento del método

El AFM es un método factorial íntimamente relacionado con el Análisis en Componentes Principales (ACP). Se diferencia de éste en que mientras en ACP las variables, de naturaleza cuantitativa, no están estructuradas en forma de partición, en el AFM sí existe una estructura de partición asociada a las variables en estudio. Sin embargo, la síntesis de información en AFM se hace por medio de un ACP (cf. Figura 1). En este contexto el AFM, además, pone en evidencia, en caso de existir, la estructura relativa de los grupos de variables. En líneas generales el procedimiento se lleva a cabo en dos etapas (ver esquema siguiente):

- **Etapa preliminar:** Consiste en realizar un ACP separado para grupo de variables inducido por la estructura de partición de las mismas. Es decir, se realiza un ACP sobre cada subtabla.
- **Etapa principal:** Consiste en realizar un ACP normado sobre la tabla global, la que se forma por yuxtaposición de todas las subtablas, previamente ponderadas por el correspondiente mayor valor propio obtenido de aplicar un método factorial clásico (ACP ó ACM) a cada una por separado. El efecto de esta ponderación se traduce en un equilibrio de los grupos.

En la etapa principal del análisis se estudian simultáneamente todas las subtablas o grupos de variables, sin perder el concepto o noción de partición definida sobre ellas. El ACP global sobre todas las variables permite un referente común que hace posible el estudio y seguimiento de la evolución o dinamismo de los distintos individuos y de las variables, proporcionando representaciones gráficas e índices de calidad que permiten el estudio de esa evolución o dinamismo. Específicamente, el AFM proporciona:

- Una representación de todas las variables en los distintos planos factoriales, tal como ocurre en cualquier ACP, con sus propias reglas de interpretación y lectura.
- Representaciones de los individuos caracterizados por los grupos de variables, lo que significa estudiar los individuos desde el punto de vista de los grupos. Esto se hace en cada nube parcial.
- Representaciones de los individuos, caracterizados por el conjunto de variables, lo que corresponde a estudiar la nube global.
- Una representación superpuesta de los grupos de variables y de los individuos. Esto corresponde a la representación y estudio en los planos factoriales de las trayectorias de los individuos,

las que se interpretan en función de su posición en el plano factorial asociado al compromiso y sobre la base de los factores que definen este compromiso. Esta representación se obtiene mediante proyección de las nubes parciales y la nube media o compromiso. La posición de los puntos parciales y el sentido de una trayectoria dan información acerca de la dirección de la evolución de un determinado individuo respecto de los factores.

- Representación de los grupos de variables (estudio de la interestructura).

Uno de los aspectos más interesantes del AFM es la representación simultánea de todas las subnubes asociadas a los grupos de variables. Esta representación se obtiene mediante proyección, y en el rol de elementos ilustrativos, de las K tablas parciales sobre los ejes factoriales asociados a la nube global.

IV. Una aplicación: Estudio de la pobreza en Chile

Quizá uno de los problemas más importantes de los países del Tercer Mundo, aunque no exclusivamente, sea la *pobreza*. Este tema, o problema, es el objetivo de políticas y medidas gubernamentales, gobierno tras gobierno, orientadas todas a superar el problema. Sin embargo, y a pesar de aquellos bien intencionados esfuerzos, el problema persiste, y lo que es peor aún, el horizonte de su erradicación definitiva muchas veces parece ser una utopía. A pesar de ello, el tema forma parte de agendas y de objetivos fundamentales de los gobiernos y de los más diversos e importantes organismos públicos, como son el Banco Mundial, el Fondo Monetario Internacional, el Programa de Naciones Unidas para el Desarrollo (PNUD), la CEPAL, entre otros.

En esta parte del trabajo se desarrolla una aplicación del AFM en orden a analizar información recogida por la encuesta CASEN de Chile desde una perspectiva estadística multivariante, para mostrar así la utilidad del método en el estudio de problemas de este tipo. Específicamente, se estudia la estructura y evolución de la pobreza a partir de algunos indicadores basados en ingresos económicos, observados en los últimos años del siglo XX.

1. El concepto de pobreza y algunos enfoques utilizados en su estudio y medición

El desarrollo y crecimiento de los pueblos no puede ni ha podido nunca desentenderse del tema de la pobreza, toda vez que a medida que se avanza en crecimiento y desarrollo, el problema se va haciendo cada vez más multidimensional, demandando implicaciones de los más diversos sectores, así como de una racionalización y conceptualización cada vez más importantes. Estos últimos procesos, concretados en estudios de sistematización e intentos de formalización del tema, se inician hacia finales del siglo XIX, y junto a ellos surgen también los esfuerzos por definir, precisar, medir, comprender, reducir y superar la pobreza, discusión y debates que siguen vigentes y lo seguirán estando hasta que no haya soluciones integrales.

Las distintas acepciones que a lo largo de tiempo se han venido utilizando ponen de relieve que el concepto de pobreza ha evolucionado y «se ha ampliado considerablemente en las últimas décadas, desde un enfoque centrado en las variables económicas de ingreso o consumo, a otras que han incorporado factores y dimensiones de la vida de las personas, tales como longevidad, analfabetismo y condiciones de salud, o bien, la vulnerabilidad, el riesgo, y la 'impotencia y la falta de voz', entre otros condicionantes que afectan significativamente a los pobres» (MIDEPLAN, 2002). Al mismo tiempo, dejan entrever un importante grado de multidimensionalidad en el tema, característica que requiere de enfoques y tratamientos pertinentes y consecuentes con esa multidimensionalidad. Como consecuencia de este natural proceso evolutivo, y dado que la mayoría de las definiciones y mediciones de la pobreza se centran en consideraciones económicas, como son el ingreso o el consumo, el tema se va transformando, poco a poco, en un tema cada vez más complejo, incluyendo otras variables y originando diversos enfoques conceptuales y medidas de pobreza. En relación a este último aspecto, sobre la medición, el estudio de la pobreza se enfoca, al menos, a los siguientes aspectos:

Incidencia: Se entiende como la cantidad de pobres en un grupo determinado de personas u hogares. Da cuenta, por lo tanto, de la magnitud de la pobreza de un país, región o zona.

Brechas: Se entienden como la comparación entre un valor dado y un parámetro de contraste, normativamente establecido. Cuando ese parámetro de contraste es una *línea de pobreza* (nivel mínimo de ingresos que permite a un individuo satisfacer sus necesidades básicas), entonces una brecha determina en qué medida los ingresos de los hogares o individuos pobres se sitúan por debajo de esa línea de pobreza.

Desigualdad: La idea de base del concepto es la 'distancia' entre ingresos extremos, y puede tratarse de ingresos de los pobres, o bien de ingresos de la población en estudio. En este último caso se habla de *desigualdad en los ingresos* o bien de *desigualdad económica*.

En relación al criterio de los ingresos existen los enfoques de la *pobreza relativa*, que se fundamenta en la idea de que las necesidades no son fisiológicamente establecidas, sino determinadas culturalmente; y el de la *pobreza absoluta*, según el cual un hogar (persona) es pobre, si sus ingresos o sus gastos agregados son inferiores a un valor equivalente al necesario para la subsistencia. Este último enfoque, además, permite el tratamiento cuantitativo de la información, en base a determinados índices, que se basan fundamentalmente en las diferencias entre hogares o personas consideradas pobres y aquellas unidades que no lo son. Entre estos índices se tienen: el **Coefficiente de GINI**, los **percentiles de ingreso**, el **Índice 20/20**, la **Inciden- cia de la pobreza**, la **Brecha promedio de la pobreza** y los **Índices FGT**. Estos últimos, desarrollados por Foster, Greer y Thorbecke, proporcionan, a su vez, medidas de descomposición de pobreza muy utilizadas en diferentes estudios (MIDEPLAN, 2002).

2. Situación en Chile a finales del siglo XX. Estudio exploratorio

En Chile, al igual que en la mayor parte de los países de América Latina, el problema reviste múltiples facetas y dimensiones, desde la política y social hasta aquella que toma en cargo la Iglesia e instituciones de beneficencia. En lo político-gubernamental, es el Ministerio de Planificación y Cooperación, MIDEPLAN, quien asume el tema en orden a definir y proponer normas y programas, con base en información objetiva, orientadas a la superación del problema, hasta donde las posibilidades del crecimiento como país lo permitan. El principal instrumento utilizado por el Estado a efectos de recoger información para, a partir de ella, elaborar y proponer propuestas de acción social, es la encuesta CA-SEN (Encuesta de Caracterización Socioeconómica

Nacional). Esta encuesta se pasa cada dos años, a partir de 1987 –con la excepción de 1989, en que fue postergada para 1990–, y recoge información en relación a la evolución de la pobreza e indigencia en el país.

En un contexto metodológico, la encuesta define al hogar como unidad primaria de análisis. Sin embargo, el tamaño del hogar puede introducir una fuente de distorsión en las medidas de pobreza utilizadas, ya que todos los hogares se consideran equivalentes en su contribución a la pobreza total, independientemente de su tamaño, pero es sabido que los hogares más pobres son frecuentemente aquéllos de mayor tamaño. Una forma de remediar esta situación es mediante la utilización del individuo como unidad básica de análisis. A pesar de esta corrección metodológica, las tendencias generales observadas al utilizar estas dos unidades de análisis no difieren significativamente, razón por la que en este trabajo se utilizaron solamente los resultados en términos de individuos.

a. Antecedentes generales y de contexto

En el plano operacional, las encuestas CASEN efectúan la medición de la pobreza en el contexto de las mediciones de pobreza absoluta, mediante el método de ingresos o método indirecto. En Chile, según este método un hogar es considerado en situación de pobreza cuando sus ingresos son insuficientes para satisfacer las necesidades básicas, alimentarias y no alimentarias, de todos sus miembros. Más específicamente, un hogar es pobre cuando su ingreso *per cápita* es inferior a 2 veces el valor de una canasta básica de alimentos, en la zona urbana, y a 1,75 veces, en la zona rural (MIDEPLAN, 2002). La línea de pobreza correspondiente al año 2000 fue de US\$ 81 en la zona urbana, y de US\$ 55 en la zona rural, aproximadamente (MIDEPLAN, 2001a, 2001b). Se asume que el costo de la canasta básica de alimentos en zonas rurales es menor al de las zonas urbanas en un 25%, supuesto que es utilizado por la CEPAL para efectuar sus estimaciones de pobreza para América Latina (CEPAL, 1996).

Aunque en los últimos años se ha observado una paulatina disminución tanto de la pobreza como de la indigencia a nivel país, gran parte de los informes también son claros, al momento de establecer sus conclusiones generales, en que esta positiva tendencia a la baja en dichos niveles, a nivel nacional, no es suficiente «para concluir que se está avanzando en todas las dimensiones que estos fenómenos presentan» (MIDEPLAN, 1998). En este sentido MIDEPLAN afirma:

En efecto, es conveniente diferenciar la incidencia de la pobreza con la intensidad de la misma. Lo que se ha observado en Chile es un continuo descenso en el porcentaje de población pobre e indigente, lo que se define como incidencia, pero cabe aún investigar qué ha ocurrido con la situación de ingresos de quienes permanecen en la pobreza e indigencia. Esta última idea se materializa en el concepto de intensidad de la pobreza que, en términos generales, señala cuál es el déficit de ingresos de los hogares pobres e indigentes en relación al ingreso necesario para satisfacer las necesidades básicas. Por otra parte, los niveles de pobreza e indigencia a nivel nacional esconden diferencias en los niveles regionales...

En este trabajo nos proponemos explorar, muy resumida e ilustrativamente, las eventuales diferencias y/o similitudes entre los distintos aspectos de pobreza y la forma en que se dan en las diferentes regiones del país, en la última década del siglo XX. La elección de este periodo no es arbitraria, sino que obedece al hecho de que es en él en donde se ha registrado, de manera sistemática y organizada, información relativa a la pobreza.

b. Los datos

Los datos utilizados en esta aplicación fueron obtenidos del sitio web del Ministerio de Planificación y Cooperación de Chile, MIDEPLAN, durante el año 2000, en la dirección www.mideplan.cl y en algunos informes elaborados por el Departamento de Estudios Sociales de la División Social de MIDEPLAN (MIDEPLAN, 1998). Específicamente, hemos utilizado la información correspondiente a:

- Coeficiente de Gini (CGI),
- Índice 20/20 (I20/20),
- Incidencia de la pobreza (INCP),
- Brecha promedio de la pobreza (BREP),
- Índice FGT2 (IFGT).

Estos indicadores, referidos a personas, no a hogares, fueron observados en las 13 regiones de Chile, y a lo largo de los años 1987, 1990, 1992, 1994, 1996 y 1998. El uso de cinco indicadores obedece a que ellos están todos presentes en todos los años estudiados.

La estructura de partición sobre las variables la define el año observado, por lo que hemos identificado cada año con un grupo de variables, en la siguiente forma:

Grupo 1: Año 1987; Grupo 2: Año 1990; Grupo 3: Año 1992; Grupo 4: Año 1994; Grupo 5: Año 1996; Grupo 6: Año 1998

Finalmente, los individuos estadísticos corresponden a las regiones político-administrativas del país vigentes al momento de construir las bases de datos. Se han numerado de 01 a 13, y nos referiremos a ellas indistintamente mediante el numeral respectivo o bien con el nombre asociado. Por ejemplo, Primera Región en lugar de 01 (la región 13 es la Región Metropolitana, en donde se ubica la Capital de la Nación).

c. Objetivos del estudio

Nos proponemos explorar la información a objeto de comprender las estructuras subyacentes a los datos, especialmente en términos de relaciones entre los distintos indicadores de pobreza y en cómo ellos pueden explicar la estructura de la misma en las distintas regiones del país, así como su dinamismo a lo largo del periodo de tiempo en estudio. En otras palabras, nos proponemos como objetivos los siguientes:

- A. Caracterizar las distintas regiones del país a partir de los diferentes indicadores de pobreza utilizados, así como su estructura intercorrelacional, mediante la exploración de la existencia de alguna tipología de las regiones de Chile según la distribución de la pobreza en cada una de ellas, y en función del tiempo estudiado.
- B. Explorar y describir la evolución de la pobreza en las diferentes regiones a lo largo del periodo estudiado.
- C. Comparar las regiones a partir de la distribución de la pobreza, sintetizada por los indicadores seleccionados, en cada uno de los periodos estudiados.
- D. Explorar y describir la estructura interna de la pobreza en las diferentes regiones, y su evolución en el tiempo.
- E. Estudiar la existencia de 'factores comunes' a los distintos periodos de tiempo y a las distintas regiones del país, como también la existencia de 'factores específicos' para alguno de los periodos o regiones estudiados.

d. Análisis y resultados

Dada la estructura de Tabla Múltiple de la información disponible, se aplicó un AFM al conjunto de datos, mediante el programa ADE4 (THIOULOUSE et al., 1997), y en lo que sigue se presentan, analizan y discuten los principales resultados obtenidos.

Los análisis individuales o parciales

El objetivo de estos análisis es proveer una visión general y *a priori* de la dimensionalidad de cada uno de los grupos de variables, para visualizar cuáles de ellos tendrán una influencia importante en la estructura comprometida. En esta parte sólo interesa la forma de la estructura más que los valores asociados a ella. En este

sentido, y en relación a los datos en estudio, puede observarse (cf. Figura 2) que todos los grupos de variables presentan una estructura claramente bidimensional. Es decir, no existen grupos casi unidimensionales que no puedan influir en muchos factores, ni grupos claramente multidimensionales que influyan en muchos factores.

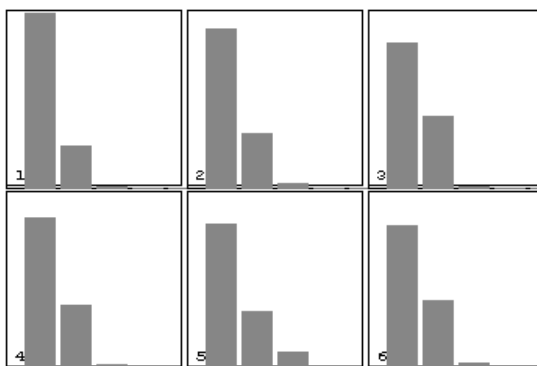


Figura 2: Descomposición de la inercia según los diferentes grupos.

Por otro lado estos análisis separados evidencian una estructura similar en cada uno de los años estudiados (cf. Figura 3, a-f). En particular evidencian que la situación de pobreza en el país se mueve, al parecer, en dos direcciones bien definidas: la *intensidad o profundidad* de la pobreza, sintetizada por las variables IFGT, INCP y BREP, y que se corresponden con el eje o fac-

tor 1. El otro factor, el segundo, está fuertemente correlacionado con las variables I20/20 y CGI, y por lo tanto resume bien el concepto de *desigualdad en la distribución de los ingresos o desigualdad económica*, concepto al que en adelante nos referiremos simplemente como *desigualdad*.

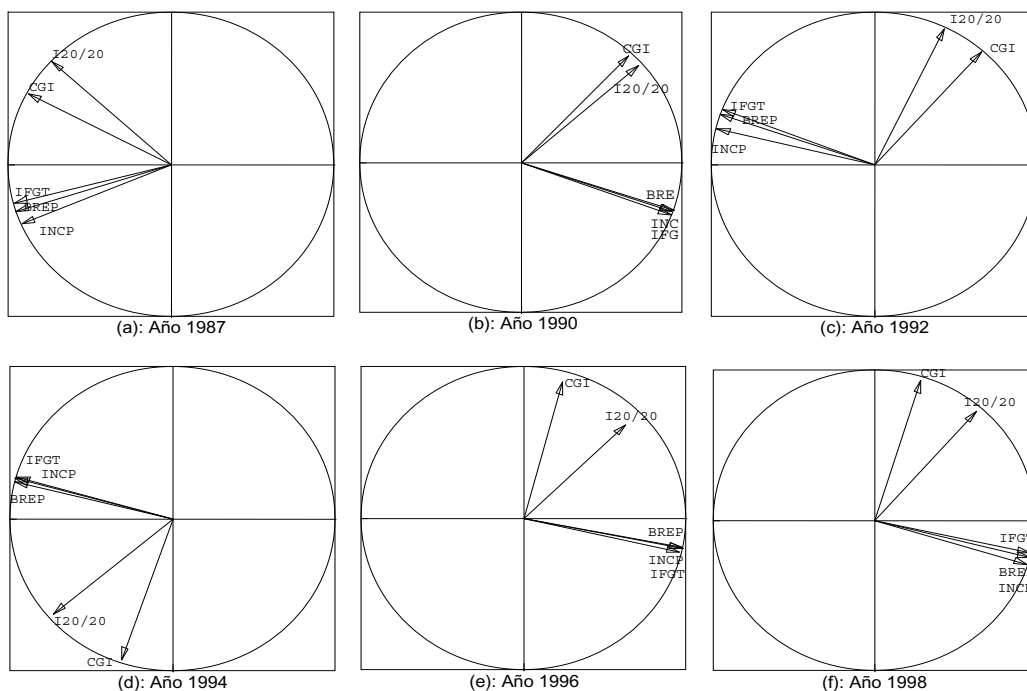


Figura 3: Configuración de las variables en estudio, en cada uno de los años considerados.

Estos elementos de síntesis se irán explorando en mayor profundidad en la medida que se avanza en los análisis y en el proceso de exploración de los datos.

El Análisis en Componentes Principales global
El ACP aplicado a la tabla ponderada, en primer lugar descompone la inercia total, que es igual a 8.699, en la forma que muestra la Tabla 1.

Tabla 1: Valores propios e inercia asociada a cada dirección del ACP global.

Num.	Eigenval.	R.Iner.	R.Sum	Num.	Eigenval.	R.Iner.	R.Sum
01	5.105	0.587	0.587	02	0.133	0.153	0.740
03	0.582	0.067	0.807	04	0.481	0.055	0.862
05	0.325	0.037	0.899	06	0.258	0.030	0.929
07	0.216	0.025	0.954	08	0.148	0.017	0.971
09	0.107	0.012	0.983	10	0.069	0.008	0.991
11	0.058	0.007	0.998	12	0.019	0.002	1.000
13	0.000	0.000	1.000				

Estos valores deben observarse en función del número de grupos de variables y de los valores propios de los análisis individuales. En el caso del primer valor propio, por tratarse de un valor (5.105) próximo al número de grupos (6) puede significar que el primer factor sea un factor común a un conjunto de grupos y por lo tanto representar una dirección de inercia importante para ese conjunto, aspecto que se irá estudiando en la medida que se disponga de más información.

Las proyecciones de las variables en el espacio determinado por este ACP global, el espacio compromiso, se posicionan en la dirección positiva de cada uno de los ejes. Globalmente las variables más relacionadas con el factor 1 son las variables INCP, BREP e IFGT, que corresponden a indicadores de *intensidad de la pobreza*. En tanto las variables más relacionadas con el factor 2 son las variables CGI e I20/20, que se asocian a indicadores de *desigualdad* (cf. Figura 4).

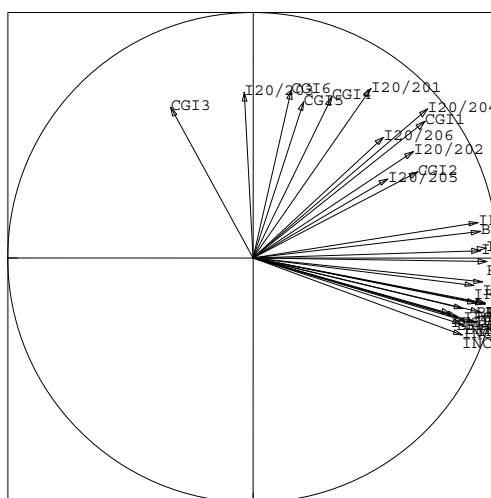


Figura 4: Representación de las variables en el espacio compromiso. El último dígito de cada símbolo indica el grupo al que pertenece la variable representada.

Por su parte las proyecciones de las regiones en el primer plano factorial del espacio compromiso (cf. Figura 5) evidencian cierta estructura que vale la pena analizar con un poco de detención. En efecto, el factor 2 pone en oposición relativa a las regiones Primera y Metropolitana (para una idea sobre la posición geográfica de estas regiones, ver Figura 1) de el grupo de las regiones que conforman la zona Centro-Norte de Chile, y de manera especial la IV región. Cabe señalar que en la Región Metropolitana se asienta la capital del país, Santiago de Chile, con un dinamismo industrial, comer-

cial, financiero y político muy particulares, como así mismo una alta densidad poblacional. La Primera Región, por su parte, se caracteriza por su actividad comercial y turística, mientras que la zona Centro-Norte se caracteriza por la actividad minera, cuyos trabajadores gozan de conocidos privilegios sociales y salariales muy diferentes de otras actividades de la región, como lo son el sector agrícola y la pesca artesanal. Estos hechos confirman al concepto *desigualdad* como parte esencial de este factor.

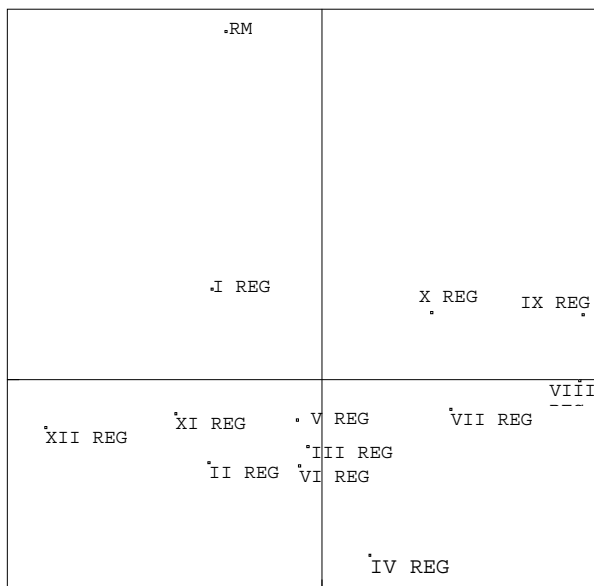


Figura 5: Representación de las regiones de Chile en el espacio compromiso.

Por otro lado, y en relación al primer eje del espacio compromiso, las regiones XI y XII aparecen en oposición de las regiones VII, VIII, IX y X, y particularmente respecto de las regiones VIII y IX, que además presentan estructuras muy similares entre sí. Es decir, la proyección de las regiones en este espacio compromiso evidencia una discriminación espacio-geométrica, respecto del factor 1, entre las regiones centro-sur, como lo son las regiones Séptima, Octava, Novena y Décima, con las regiones XI y XII, las más australes del país. Estas últimas son regiones que mantienen importantes problemas de comunicación con el resto del territorio nacional, y cuya actividad económico-productiva descansa en la industria del turismo, ganadería, pesca artesanal y del petróleo (esta última actividad es propia de la XII Región). Son regiones no muy expues-

tas a los estímulos consumistas como lo están las otras regiones del país. Las regiones VIII y IX, en cambio, son las regiones más pobres del país, y muy particularmente la IX, con un porcentaje muy alto de campesinos y mapuche. Su principal actividad económico-productiva se centra, en el caso de la Octava Región, en la industria siderúrgica y petroquímica, pesquera y sobre todo forestal. La IX Región, por su parte, es una región eminentemente agrícola y forestal. Se confirma también así la apreciación anterior en el sentido que este factor sintetiza bien el concepto de *intensidad de la pobreza*.

En resumen, podemos apreciar, desde una perspectiva global, que en relación a la pobreza las regiones de Chile pueden ordenarse según dos factores importantes: Un primer factor, que denominamos *intensidad de*

la pobreza, y al que se vinculan principalmente las regiones de la zona centro-sur de Chile y las de la zona austral; y un segundo factor, que resume el concepto de *desigualdad*, vinculado más bien a las regiones de la zona norte (con la excepción de la primera región, que es esencialmente turística y comercial) con los más bajos índices de pobreza. En este contexto, las regiones Metropolitana y IV se presentan, respectivamente, como las regiones con mayor y menor desigualdad, respectivamente; mientras que la XII región aparece como la región menos pobre del país, y las regiones VIII y IX como las más pobres a nivel nacional.

Por otro lado, la similitud entre las proyecciones de las nubes sobre un determinado eje se mide por medio de

las correlaciones entre cada una de las nubes y la proyección de la nube global en el espacio compromiso. En nuestro caso, estas correlaciones están contenidas en la Tabla 2, y sobre la base de sus elevados valores se puede concluir que el primer factor es un factor común a los seis grupos de variables, tal como se adelantara en comentarios anteriores. Lo anterior significa que el concepto de *intensidad de la pobreza* es un tema transversal a los seis periodos considerados en el estudio. No ocurre lo mismo con el factor *desigualdad*, que se muestra como común sólo para las variables del año 1994.

Tabla 2: Correlaciones entre factores y grupos

	fac 1	fac 2
Tab 1	0.919	0.580
Tab 2	0.905	0.626
Tab 3	0.905	0.624
Tab 4	0.911	0.843
Tab 5	0.957	0.682
Tab 6	0.963	0.652

A pesar de la característica común global del factor 1, la descomposición de la inercia de las dos primeras componentes del ACP global, o equivalentemente, la inercia proyectada de cada tabla sobre el subespacio generado por los dos primeros vectores propios del ACP (cf. Tabla 3), pone en evidencia la importancia relativa del factor 1 en los periodos 1996 y 1998, y la menor

importancia de este factor en los otros periodos, particularmente en el periodo 1992. Aunque en esta tabla también se puede evaluar la calidad de representación de cada tabla en el espacio compromiso, se debe tener presente que esta calidad de representación es siempre débil (ESCOFIER & PAGÈS, 1992).

Tabla 3: Inercia proyectada de los grupos de variables

	Factor1	Factor2	Inercia total		Calidad de representación
			Proyectada	de la nube	
Grupo 1: Año 1987	0.834	0.206	1.039	5	20.8%
Grupo 2: Año 1990	0.818	0.136	0.953	5	19.1%
Grupo 3: Año 1992	0.787	0.287	1.073	5	21.5%
Grupo 4: Año 1994	0.829	0.295	1.123	5	22.5%
Grupo 5: Año 1996	0.915	0.191	1.106	5	22.1%
Grupo 6: Año 1998	0.924	0.220	1.143	5	22.9%
Total	5.105	1.333	6.438		

La razón Inercia inter/Inercia total, en relación al factor 1, es igual al 79.30%, que viene a confirmar el carácter de común de este factor. Este resultado se traduce en que los puntos que representan a una misma región a través de los diferentes grupos de variables (periodos) están globalmente próximos entre sí (cf. Figura 8). Por otra parte, las proyecciones de los distintos grupos sobre el plano factorial del espacio compromiso (cf. Fi-

gura 6) evidencian una fuerte asociación entre todos los grupos con el factor 1, *intensidad de la pobreza*; y de los años 1992 y 1994 con el factor 2, *desigualdad*. Esto significa que los años 1992 y 1994 son periodos asociados más bien a la *desigualdad* que a la *intensidad de la pobreza*, mientras que los otros periodos parecen definirse por cuestiones relativas a este último concepto.

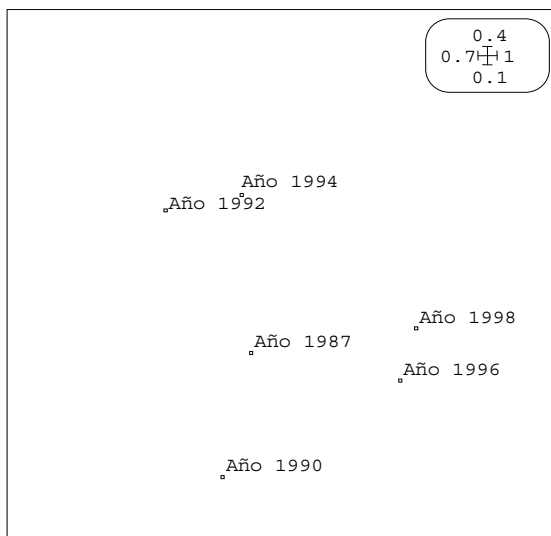


Figura 6: Representación de los grupos de variables

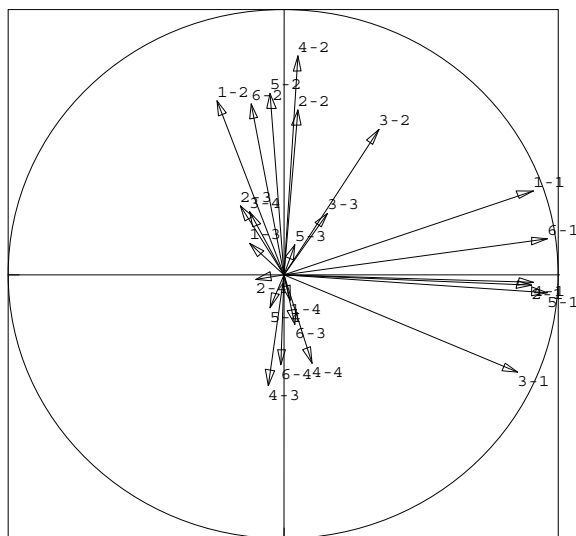


Figura 7: Correlaciones entre primeras componentes de cada grupo y las del AFM

Equivalentemente a lo anterior, el gráfico confirma que el primer factor es una dirección de inercia importante para los grupos correspondientes a los años 1987, 1990, 1996 y 1998. El factor *desigualdad*, por su parte, es claramente una dirección de inercia importante para los periodos 1992 y 1994. Por último en esta representación se pone de manifiesto la similitud de estructuras entre los periodos 1996 y 1998, por una parte, y 1992 y 1994 por otra. Retomando su vinculación con los factores en estudio, puede afirmarse que la estructura de la *intensidad de la pobreza* en el año 1996 es muy similar a la del año 1998; y que el tema de la *desigualdad económica* en los años 1992 y 1994 se presentó de forma similar.

Hemos establecido y afirmado, casi reiteradamente, que el primer factor es una dirección de inercia importante para cada uno de los seis periodos en estudio. La Figura 7² nos permite ir un poco más allá: este primer factor común es un factor próximo a la dirección principal de inercia de los seis grupos. Dicho de otra forma, está fuertemente relacionado con la primera y más importante dirección de inercia de cada grupo, como lo es la primera componente principal. Con la posible excepción del año 1992 (grupo 3), la segunda componente principal del AFM está fuertemente relacionada con la segunda dirección de inercia de cada grupo o periodo.

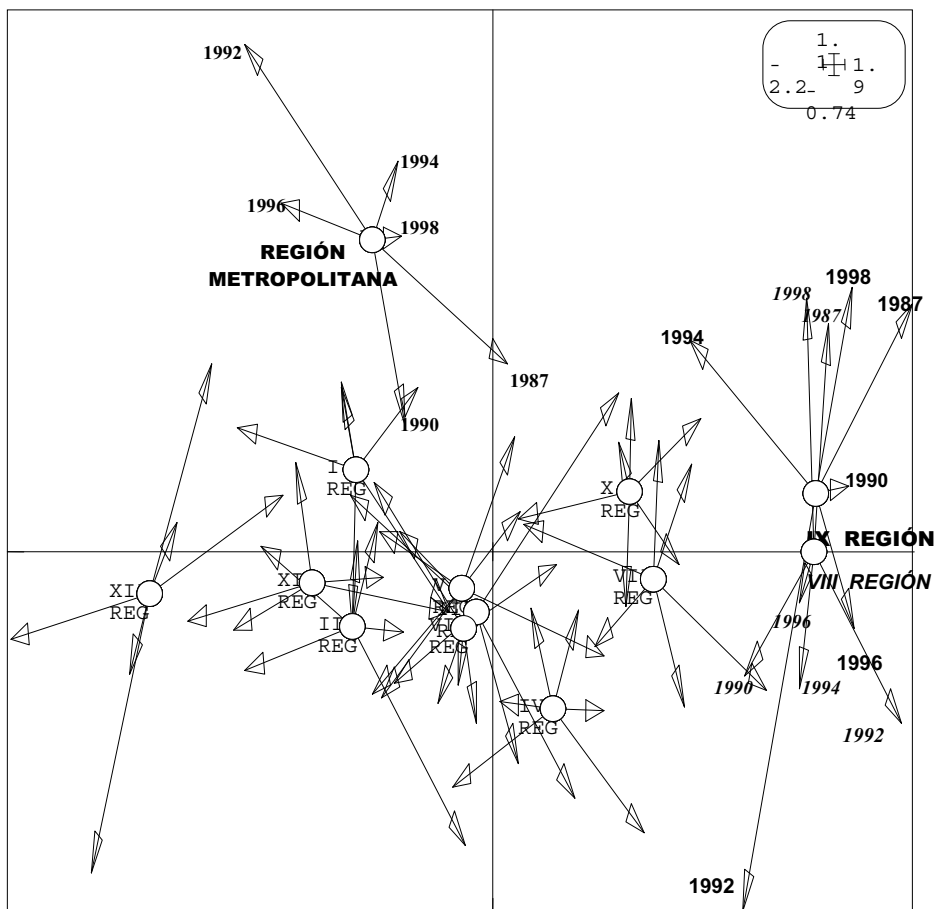


Figura 8: Representación de grupos y regiones

Para finalizar esta serie de lecturas e interpretaciones, demos ahora una mirada a la figura 8. En ella están proyectadas las distintas regiones del país según los periodos en estudio. Específicamente, cada punto–región se corresponde con el baricentro de los seis grupos que lo describen desde el punto de vista de las variables que conforman cada grupo. Por ahora comentaremos brevemente la situación de las regiones VIII, IX y Metropolitana, conscientes de que ese análisis debe extenderse a las otras 10 regiones restantes.

Ya hemos señalado, en el contexto del análisis de la figura 5, el grado de similitud entre las regiones Octava y Novena y sus principales características socio-productivas, como también hemos llamado la atención sobre el hecho que la Región Metropolitana aparezca «alejada» del resto de las regiones. Esta última región se ha presentado como un caso correlacionado con el factor *desigualdad*, y de manera especial el grupo asociado al año 1994. Es decir, la componente principal de

la pobreza en general en esta región es la desigualdad social, expresada como una forma más de pobreza. Se observa, hacia finales de periodo, una tendencia hacia estructuras similares a las del comienzo del periodo, acompañada de una mejora en los niveles del factor *intensidad de la pobreza* en estos últimos años.

En relación a las regiones Octava y Novena, por su parte, el factor *intensidad de la pobreza*, tal como ha sido definido antes, parece ser de tipo estructural, pues en la dirección de este factor ambas regiones presentan un grado muy bajo de variabilidad, con la excepción, quizá, del año 1992 en la Novena Región, al que además se asocian los más bajos índices de este factor. El «dinamismo» de la pobreza en estas regiones parece definirse más bien por el factor de *desigualdad de ingresos*. En este contexto ambas regiones han experimentado situaciones de desigualdad muy similares en los años 1987 y 1998. Además hacia finales del periodo en estudio el factor de la desigualdad parece

acentuarse drásticamente, llegando a niveles similares a los que había en el año 1987. Las causas de este fenómeno no son conocidas, pero bien valdría la pena estudios más profundos, toda vez que, como hemos adelantado, se trata de las regiones más pobres del país. También puede ser interesante estudiar las condiciones que, en algunos periodos, llevaron a estas regiones a sus más bajos índices de desigualdad social, como es el año 1992 en la Novena Región y los años 1990, 1992 y 1994 en la Octava Región.

A manera de resumen, el AFM aplicado a esta base de datos ha permitido caracterizar la pobreza en Chile, sobre la base de estos datos, en el periodo estudiado, posibilitando análisis desagregados en función de las distintas regiones del país, y poniendo de manifiesto algunas características que mediante otra técnica habrían resultado más difíciles de explicitar.

Comentarios finales

La aplicación del Análisis Factorial Múltiple a la información disponible sobre indicadores de pobreza en Chile no sólo ha confirmado una estructura general *a priori* dada a conocer por diferentes organizaciones civiles, tanto del país como internacionales –MIDEPLAN, Fundación para la Superación de la Pobreza, Fundación Chile Solidario, PNUD, CEPAL, BID, Banco Mundial, FMI, entre otras–, sino que ha permitido también explorar con cierto grado de profundidad en sus expresiones más particulares y específicas, especialmente explorando y describiendo relaciones e interrelaciones entre los índices utilizados, las regiones y los periodos de tiempo considerados. Con base en ese proceso exploratorio/descriptivo, podemos concluir:

- 1) A grandes rasgos se ha establecido la existencia de dos factores importantes a la hora de explicar la pobreza: El eje de la *intensidad de la pobreza*, caracterizado por la Incidencia de la pobreza, Brecha de la pobreza y el índice FTG2, de cada grupo. El segundo factor puesto de manifiesto por los datos es el de la *desigualdad en la distribución de los ingresos*, definido básicamente por indicadores como el Coeficiente de GINI y el Índice 20/20.
- 2) El primer factor se trata de un factor presente a lo largo de todo el periodo estudiado, indicando con ello que la *intensidad de la pobreza* es un problema al que las distintas administraciones y políticas gubernamentales no han podido dar solución.

- 3) Las regiones Octava y Novena son las regiones con los más altos índices en el factor de la *intensidad de la pobreza*, seguidas de las regiones 7ª y 10ª, todas regiones de la zona centro-sur de Chile, en oposición a las regiones más australes, que presentan los más bajos índices de este factor. Destaca en este último grupo la XII región, la región más austral de Chile, de la que ya hemos hablado más arriba, pero que respecto de ella es necesario agregar que se trata, a la luz de los análisis, de una de las regiones en donde la pobreza es 'menos notoria' que en el resto del país.
- 4) Respecto del segundo factor, de la *desigualdad en la distribución de los ingresos*, se ha concluido que es en las regiones Metropolitana y Primera, pero especialmente en la Metropolitana, en donde se dan los más altos valores, mientras que los más bajos índices de *desigualdad* se encuentran en las regiones de la zona centro-norte, y muy particularmente en la IV Región. No obstante, es necesario precisar que este factor es característico más bien de los años 1992 y 1994.
- 5) Existen ciertas interrogantes relativas a la estructura y dinamismo de la pobreza en las regiones más arriba comentadas, tales como ¿qué hechos o circunstancias llevan a las regiones Octava y Novena a una estructura similar de la pobreza en los inicios del periodo como hacia finales del mismo? También es pertinente preguntarse por las circunstancias que llevaron a la Novena Región, en el periodo 1990–1992 a presentar los más bajos índices de desigualdad, incluso con leve mejoramiento de sus niveles de *intensidad de la pobreza*.

En este sencillo esfuerzo por comprender uno de los temas más preocupantes a nivel país, como es la pobreza en todas sus formas y expresiones, el AFM ha permitido poner de relieve no sólo aspectos y relaciones *a priori* conocidas, sino que también ha sido capaz de revelar importantes inquietudes, como el carácter aparentemente cíclico de la *desigualdad* y de la *intensidad de la pobreza* en las regiones Metropolitana, Octava y Novena –que son las regiones que aquí hemos estudiado, y por lo tanto tales resultados no pueden asociarse a las demás regiones del país–, hecho que sin duda justifica la necesidad de mayores estudios y en mayor profundidad, puesto que lo que corresponde esperar es que la estructura de pobreza de los países avance hacia estados de bienestar, sin 'retroceder' a aquellos que han sido superados.

La situación de pobreza de muchos pueblos, y en particular la del caso que nos ocupa, interpela a la sociedad en su conjunto, dígase especialmente actores políticos y sociales, por una comprensión profunda del problema, para que, sobre bases cada vez más objetivas, se construya un horizonte de esperanza hacia el que se camine por los senderos de la solidaridad, de la justicia social y del bien común, por medio de un mejoramiento sustantivo de la situación en la que viven miles de hogares chilenos.

THIOULOUSE et al. 1997. *The Ade4 package*. <http://pbil.univ-lyon.fr/ADE4/>.

Notas

¹ Investigación financiada por el Proyecto MECESUP 0007 UCT, Universidad Católica de Temuco.

² En ese gráfico, el símbolo $a - b$ indica la componente b del grupo a .

Bibliografía

CARROLL, J.D and CHANG, J.J. 1970. «Analysis of individual differences in multidimensional scaling via an N-way generalization of «Eckart-Young» decomposition». *Psychometrika*, 35: 283-320.

CARROLL, J.D. 1968. «A generalized of canonical correlation analysis to three or more sets of variables». Proceedings of 76th annual convention of the *American Psychological Associations*, 227-228.

CEPAL. 1996. «La Sensibilidad del Indicador de Pobreza. Un Análisis a Partir de Diferentes Opciones Metodológicas».

ESCOFIER, B. & PAGÈS, J. 1992. *Análisis factoriales simples y múltiples. Objetivos, métodos e interpretación*. (Trad.). Servicio Editorial Universidad del País Vasco.

HOTELLING, H. 1936. «Relation between two sets of variables». *Biometrika*, 28, pp 129-149.

KETTENRING, J.R. 1971. «Canonical analysis of several set of variables». *Biometrika*, 58: 433-460.

LE DIEN & PAGES. 2003. «Analyse factorielle multiple hiérarchique». *Revue de Statistique Appliquée*, 51 No. 2. pp. 47-73.

MIDEPLAN. 1998. *Evolución de la Pobreza e Indigencia en Chile, 1987-1996*. 113 págs.

_____ (2001a). *Pobreza, desigualdad y convergencia regional: Un análisis para el periodo 1990-1998*. Unidad de Estudios Prospectivos. 109 págs.

_____ (2001b). *Pobreza e indigencia e impacto del gasto social en la calidad de vida*. Informe Ejecutivo. 28 págs.

_____ (2002). *Síntesis de los principales enfoques, métodos y estrategias para la superación de la pobreza*. Documento de trabajo. 44 págs.