

EI INDICADOR MULTIDIMENSIONAL DE CONDICIONES DE VIDA

Elkin Castaño

Abstract

The objective of this document is to show as the techniques of the variable statistic analysis can be used to construct the Multivariable Index of the Poverty. The author raises that those are the data say which of the selected variables are the appropriate in each dimension of the poverty and which are their optimal considerations. The idea that the poverty can be measured using a great number of variables has increasing acceptance on the part of the social investigators, because the traditional methodologies due to their unidimensionalidad or that the variables no longer mark has seen the necessity to propose different multidimensional indicators which try to capture the different facets from the poverty.

Resumen.

El objetivo de este documento es mostrar como las técnicas del análisis estadístico multivariado pueden ser empleadas para construir el Índice Multivariado de la Pobreza. El autor plantea que sean los datos mismos los que digan cuáles de las variables seleccionadas son relevantes en cada dimensión de la pobreza y cuáles son sus ponderaciones óptimas. La idea de que la pobreza puede ser medida utilizando un gran número de variables tiene cada vez mayor aceptación por parte de los investigadores sociales, debido a que las metodologías tradicionales debido a su unidimensionalidad o que las variables ya no marcan se ha visto la necesidad de proponer distintos indicadores multidimensionales los cuales tratan de capturar las diferentes facetas de la pobreza.

1. Introducción

La idea de que la pobreza puede ser medida utilizando un gran número de variables tiene cada vez mayor aceptación por parte de los investigadores

sociales. Durante mucho tiempo, y específicamente desde la introducción del concepto económico de la pobreza debido a Booth (1892) y a Rowntree (1901), el indicador de referencia utilizado ha sido el ingreso o el gasto per-cápita. Pero mientras estos indicadores son medidas razonablemente precisas y útiles del comportamiento económico, distintos autores entre los cuales se encuentran Townsend (1993), Ravallion (1996) y Tsui (2002), le han hecho duras críticas debido a su unidimensionalidad y se han propuesto distintos indicadores multidimensionales los cuales tratan de capturar las diferentes facetas de la pobreza. Desde el trabajo de Townsend (1979), se ha venido reconociendo que otros aspectos de la vida humana, no necesariamente relacionados con el ingreso, impactan el desarrollo humano. Estos incluyen el acceso a los bienes públicos, la salud, la educación, las condiciones de la vivienda, la satisfacción de la vida, etc. Además, la evolución del pensamiento conceptual sobre la pobreza hacia las funcionalidades y capacidades de los individuos, iniciada por Sen (1993), también ha incrementado el interés por medidas de pobreza multidimensionales.

La consecuencia de esta revolución conceptual es la noción ampliada de pobreza la cual incluye vulnerabilidad, exposición al riesgo, el empoderamiento (World Bank, 2001). Hoy en día, la pobreza no se refiere solamente a la falta de habilidad de los individuos u hogares para obtener los recursos suficientes para satisfacer sus necesidades básicas (Townsend, 1993), ni se considera una simple dimensión económica y monetaria. Cada vez más se entiende como una privación en varios dominios de la vida humana. Esta privación desde la perspectiva multidimensional, incluye tanto mediciones de variables cuantitativas como cualitativas.

Desde el punto de vista empírico, los últimos años han sido testigos de una búsqueda intensa de aproximaciones adecuadas para medir la pobreza multidimensional. Estas aproximaciones incluyen la exclusión social de Rene Lenoir (1974), la aproximación axiomática multidimensional y el índice de desarrollo humano (IDH) del Programa de las Naciones Unidas para el

Desarrollo, PNUD, (1997). Éste último combina esperanza de vida, educación y salud. Aunque ampliamente usado, este índice cada vez tiene más críticos por que no considera la dimensión del ingreso y por atribuir arbitrariamente ponderaciones iguales a cada dimensión. También, la elección de las variables elegidas es algo arbitraria y puede no reflejar las preferencias y realidades de las personas del país bajo estudio (Booyesen, 2002).

La aproximación axiomática multidimensional comienza con la especificación de una función general de la forma $P(x, z) = F[\pi(x_i, z)]$, donde F y $\pi(\cdot)$ están basadas en algunos axiomas que estipulan como pueden ser calculados los indicadores de pobreza (Bourguignon and Chakravarty, 2002 and Bibi, 2005). Sin embargo, la forma funcional de la ecuación es bastante arbitraria y subjetiva.

De acuerdo a la literatura más reciente, la pobreza está conceptualizada en términos de exclusión de la vida de sociedad debido a la falta de recursos, mientras que la exclusión significa la experimentación de varias formas de lo que la sociedad considera como una privación grave. En consecuencia, la pobreza sería mejor tratada cuando los indicadores multidimensionales y no monetarios son complementados con los monetarios para identificar mejor a los pobres.

A pesar de estas ventajas, las medidas de pobreza que incorporan información de muchas variables también tienen sus problemas, principalmente el de enfrentarse a la multidimensionalidad y el uso de variables no-monetarias. Cuando se trata de hacer operacional un concepto de pobreza multidimensional, se deben enfrentar muchos desafíos. Puesto que todas estas elecciones afectan en forma significativa el índice multidimensional resultante, es importante clarificarlas.

El objetivo de este documento es mostrar como las técnicas del análisis estadístico multivariado pueden ser empleadas para construir el IMP, en el

sentido de que sean los datos mismos los que digan cuáles de las variables seleccionadas son relevantes en cada dimensión de la pobreza y cuáles son sus ponderaciones óptimas, en lugar de hacer suposiciones subjetivas a priori.

2. Índices Compuestos

Cuando la pobreza es conceptualizada como una construcción multidimensional, debería ser medida a través de la agregación de diferentes variables que indiquen la privación experimentada por los individuos. De acuerdo a esto, la medición de la pobreza multidimensional generalmente tiene que ver con la incorporación de la información dada por las variables en el índice compuesto de pobreza. El procedimiento general en la estimación de índices compuestos tiene en cuenta siguientes etapas:

- La escogencia de las variables que se van a considerar
- La definición de un esquema de ponderación para cada variable
- La agregación de las variables
- La identificación de un umbral que separa los individuos pobres y los no pobres

Estas consideraciones deben ser bien entendidas en la construcción del índice.

El primer paso para la construcción de una medida resumen de pobreza tiene que ver con la selección adecuada de los indicadores. Esto depende de la disponibilidad de la información, pero hay que tener en cuenta que las variables consideradas afectan el índice resultante. La selección de las variables básicas se basa en la elección arbitraria de los investigadores quienes deben enfrentar el trade-off entre posibles redundancias causadas por la sobreposición de información y el riesgo de perder información. Una solución parcial a tal arbitrariedad está dada mediante el uso de técnicas estadísticas multivariadas, las cuales permiten revelar la existencia de correlaciones (o, en

general, dependencias) entre las variables básicas y retener solamente un subconjunto que mejor resuma la información disponible.

Una vez un conjunto preliminar de variables se ha seleccionado, su agregación en un índice compuesto implica la escogencia de una estructura adecuada de ponderaciones. Para esto, en la literatura se han propuesto diferentes técnicas de obtener dichos pesos. En primer lugar, algunos estudios aplican ponderaciones iguales a las variables (Townsend, 1979; PNUD, 1997 y Nolan y Whelan, 1996), y de esta forma evitan la necesidad de asignar importancia diferente a las distintas dimensiones. En segundo lugar, en la construcción de índices de pobreza compuestos, las variables se combinan usando ponderaciones determinadas por medio de un proceso de consultas con expertos en pobreza y analistas de política. Aunque este enfoque es una mejora a la primera solución, aún involucra decisiones subjetivas sobre el nivel de bienestar aportado por cada variable. Tercero, las ponderaciones pueden ser aplicadas de forma que reflejen la calidad de la información de las variables, dando menos peso a aquellas variables donde hay problemas con los datos o hay un gran número de datos faltantes (Rowena et al. 2004). La confiabilidad del índice compuesto de pobreza puede ser mejorada si se le da más peso a las variables con mejor calidad de datos. Sin embargo, este método puede dar más peso a variables que son más fáciles y rápidamente de medir en lugar de asignarlo a variables que indiquen mayor bienestar, las cuales pueden ser más problemáticas. Cuarto, un método más objetivo es imponer un conjunto de ponderaciones usando los precios de las variables. Sin embargo, esto solamente es posible si los precios están disponibles para todos los bienes y servicios. Desafortunadamente este no es el caso. Generalmente los encuestados son incapaces de valorar sus bienes realísticamente y sus respuestas probablemente contendrán grandes errores. Finalmente, otros estudios desarrollan índices compuestos ponderando las variables sobre la base de sus frecuencias relativas, o basados en métodos estadísticos multivariados tales como los del análisis estadístico factorial. Este enfoque es el propuesto y será discutido en la siguiente sección.

Por último, la identificación de los hogares o individuos pobres requiere de la definición de un umbral y éste es un tema que presenta varios problemas teóricos y empíricos. Independientemente de una elección particular del umbral, la identificación de quienes son considerados pobres siempre tiene un grado de arbitrariedad.

3. Aproximaciones usando técnicas del análisis estadístico multivariado

Como se mencionó anteriormente, existen diferentes formas de asignar las ponderaciones de las variables para la construcción de un índice de pobreza multidimensional. La aproximación estadística proporciona diferentes alternativas para seleccionar y agregar variables en forma de índices sin necesidad de hacer supuestos a priori en el esquema de ponderación de las variables. A continuación, se presentan solamente las características relevantes de los métodos estadísticos multivariados basados en el análisis factorial que son útiles para la construcción de un índice de pobreza multidimensional. Para una descripción más completa de los procedimientos expuestos véase Johnson y Wichern (1998), Mardia, Kent y Bibby (1979) y Gifi (1990) entre otros.

3.1. El análisis de Componentes Principales: el caso de variables numéricas

Uno de los documentos más influencias que usan el Análisis de Componentes Principales (PCA) para la construcción de índices socio-económicos en desarrollo económico y estudios de población es el de Filmer & Pritchett (2001). Estos autores emplearon datos sobre los hogares tales como bienes durables, tipos de acceso a servicios higiénicos, número de cuartos de la vivienda, y materiales de construcción, para construir un índice socio-económico usando el PCA.

El interés del Análisis de Componentes Principales (el cual forma parte de los procedimientos estadísticos del Análisis Factorial) es explicar la estructura de varianzas y covarianzas de un conjunto de p variables por medio de unas pocas combinaciones lineales de ellas. En general, sus objetivos son dos: Reducción de la dimensión de los datos multivariados y la interpretación del problema en la nueva dimensión.

Aunque se necesitan p combinaciones lineales para reproducir la variabilidad total del sistema, con mucha frecuencia, la mayor parte de esta variabilidad puede ser explicada por medio de un pequeño grupo de k componentes principales ($k < p$). Si esto es posible, entonces hay tanta información en unas pocas componentes principales como en el sistema de variables originales. Con frecuencia, un análisis usando componentes principales revela relaciones que antes nos se sospechaban y por consiguiente permite interpretaciones más potentes de los datos. En nuestro contexto de la construcción de indicadores de pobreza, el PCA ha sido aplicado, entre otros autores, por Klasen (2000), Nagar y Basu (2001) y Kolenikov y Angeles (2004).

La aplicabilidad de las técnicas de análisis factorial está generalmente limitada por la clase de datos disponible. Específicamente, en principio el PCA puede ser aplicado si todas las variables son numéricas (es decir, las variables son cuantitativas o continuas) y las relaciones entre las variables son lineales (Gifi, 1990). Sin embargo, las variables usadas en el análisis de pobreza pueden ser tanto numéricas como categóricas, medidas en nivel nominal u ordinal. En consecuencia, el PCA no es el método más apropiado para estos casos. El problema aparece debido a que las variables cualitativas no tienen un origen o una unidad de medida y por tanto, las medias, varianzas y covarianzas de estas variables no tienen interpretación real. Como el PCA se basa en la estimación de la matriz de varianzas y covarianzas (o en la matriz de correlación), el PCA estándar no es adecuado. Para otras dificultades véase (Kamanou, 2005).

El modelo de Componentes Principales y el esquema de ponderaciones

Si X_1, X_2, \dots, X_p es el conjunto de características cuantitativas que queremos resumir, entonces el PCA proporciona las p nuevas variables:

$$Y_1 = a_{11} X_1 + a_{12} X_2 + \dots + a_{1p} X_p \quad (\text{Primera Componente})$$

$$Y_2 = a_{21} X_1 + a_{22} X_2 + \dots + a_{2p} X_p \quad (\text{Segunda Componente})$$

.....

$$Y_p = a_{p1} X_1 + a_{p2} X_2 + \dots + a_{pp} X_p \quad (\text{P-ésima Componente})$$

Con las siguientes propiedades (bajo ciertas condiciones):

- ✓ Las componentes son combinaciones lineales de las variables originales, donde a_{ij} es la ponderación (peso) que tiene la variable X_j sobre la componente i .
- ✓ Las componentes son resúmenes de la información de las variables originales.
- ✓ La primera componente principal Y_1 contiene la mayor cantidad de información de las variables originales X_1, X_2, \dots, X_p .
- ✓ La segunda componente principal Y_2 contiene la mayor cantidad de información después de la primera componente, y así sucesivamente.
- ✓ La información de la primera componente principal es única, en el sentido de que no la comparte con las demás componentes.
- ✓ La información de la segunda componente principal es única, en el sentido de que no la comparte con las demás componentes, y así sucesivamente.

De esta manera, la metodología del PCA es útil para el propósito de construir el indicador de pobreza multidimensional, pues permite obtener el indicador como la

combinación lineal que contiene **máxima información** de las variables que lo componen y su información **es única (no es compartida)** por las otras combinaciones lineales).

Sin embargo, generalmente las variables seleccionadas presentan un nivel mixto de medición: se utilizan variables cualitativas y cuantitativas. Para este conjunto de variables con nivel de medición mixto, no podemos aplicar directamente el PCA tradicional con el fin de obtener el indicador, pues esa técnica opera sobre variables cuantitativas. La estrategia para poder usar la técnica del PCA consiste en volver todas las variables cuantitativas a través de la técnica del “Optimal Scaling” o Cuantificación Óptima. Este procedimiento asigna valores numéricos a las categorías de las variables de forma tal que se maximice la varianza (se maximice la cantidad de información) de la primera componente principal teniendo en cuenta las características de medición de los datos. Una vez cuantificadas las categorías de las podemos emplear legítimamente el modelo de Componentes Principales para obtener directamente el indicador o podemos usar las variables cuantificadas para hacer otro tipo de análisis multivariados tradicionales. (Factores, Cluster, etc.)

El procedimiento se encuentra descrito en Young (1981), Kuhfeld, Sarle, y Young, (1985), Saporta, (1983), Young, Takane, y de Leeuw, J. (1978, 1985), Van de Geer, (1993). Algunas aplicaciones a indicadores de condiciones de vida se encuentran en Castaño y Moreno (1994), Sarmiento et al (1996), Castaño, Correa y Salazar (1998), Castaño y Valencia (1999^a), Castaño (1999^b) y Castaño (2000). Algunas propiedades del indicador han sido estudiadas en Cortés et al (1998) y Castaño (1999).

3.2. Metodología general para la construcción del indicador multidimensional

Las etapas para la construcción del indicador son las siguientes:

i) Para las variables de cada dimensión (observadas en diferentes niveles de medición) se usa el método de cuantificación óptima. El producto de esta etapa es obtener las transformaciones óptimas de las variables cuantitativas y cualitativas simultáneamente.

ii) Sobre las variables transformadas, se emplea el Análisis de Factores para identificar agrupamientos o factores en la dimensión estudiada. Para cada factor identificado se usa el PCA para construir un indicador.

iii) Los indicadores para cada dimensión son resumidos en un único indicador de la dimensión a través del PCA.

iv) Una vez obtenidos los indicadores de todas las dimensiones, se usa de nuevo el PCA para construir un indicador global de pobreza .

El Apéndice A presenta un diagrama de flujo para la construcción del indicador en una dimensión.

4. Dimensiones del indicador de calidad de vida

Las siguientes son las dimensiones que se considerarán para la elaboración de un indicador de condiciones de vida multidimensional: a) Vulnerabilidad, b) acceso al conocimiento, c) acceso al trabajo , d) escasez de recursos, e) desarrollo infantil, f) carencias habitacionales, g) participación social, h) medio ambiente, i) cultura y recreación, j)movilidad, k) seguridad y l) autopercepción de la calidad de vida.

Cada una de las dimensiones mencionadas, requerirá de diferentes indicadores que permita representarla adecuadamente.

Referencias Bibliográficas

Bibi, S. (2005) *Measuring Poverty in a Multidimensional Perspective: A Review of the Literature*, Working paper PMMA 2005-07.

Booth, C. (1892) *The Inhabitants of Tower Hamlets their condition and occupations*, Journal of Royal Statistical Society, vol. 50, pp 326-340.

Booyesen, F. (2002) *An Overview and Evaluation of Composite Indices of Development*, Social Indicators Research, Vol. 59, No. 2, pp. 115-151.

Bourguignon, F. and Chakravarty, S. R., (2002) *Multidimensional Poverty Orderings*, Mimeo.

Castaño, E., y Moreno, H. (1994), Metodología Estadística para la Selección de Variables del Sistema de Beneficiarios de Programas Sociales, SISBEN, Misión Social, DNP.

Castaño, E., Correa, C. y Salazar B. (1998), “La construcción de un indicador de Calidad de Vida para la ciudad de Medellín”, mimen DNP, Misión Social.

Castaño, E. y Valencia, L.A. (1999^a), Indicador de Calidad de los Estratos para el Área Metropolitana de Medellín, Revista Lecturas de Economía, Vol. 50, 1998.

Castaño, E. (1999) Diseño del sistema de identificación de beneficiarios de programas sociales, Informe Técnico, CONAM, Quito.

Castaño, E. (2000) Construcción de un indicador de Calidad de vida para Jamaica, BID, 2000.

Filmer, D. & Pritchett, L. (2001), ‘Estimating wealth effect without expenditure data— or tears: An application to educational enrollments in states of India’, *Demography* **38**, 115–132.

Gifi, A. (1990) *Nonlinear Multivariate Analysis*, in Heiser W., Meulman J. J., van den Berg, G. (Eds), J. Wiley and Sons, Chichester

Johnson, R. A., Wichern, D. W. (1998), Applied multivariate statistical analysis. Prentice Hall, New Jersey.

Kamanou, G. (2005) *Methods of Factor Analysis for Ordinal Categorical Data and Application to Poverty Data*; International Conference, The Many Dimensions of Poverty, Brazil

Klasen, S. (2000) *Measuring poverty and deprivation in South Africa*. Review of Income and Wealth, 46, 33–58.

Kolenikov, S. and Angeles, G. (2004) *The Use of Discrete Data in PCA: Theory, Simulations, and Applications to Socioeconomic Indices*, Working Paper of Measure/Evaluation Project, No. WP-04-85 Carolina Population Centre, University of North Carolina.

Krishnakumar, J. and Nagar, A.L. (2007) *On Exact Statistical Properties of Multidimensional Indices Based on Principal Components, Factor Analysis, MIMIC and Structural Equation Models*, Social Indicators Resources, Springer Science and Business media.

Mardia, K.V., Kent, J.T. y Bibby, J.T. (1979), *Multivariate Analysis*, Academic Press, New York.

Nagar, A. L., and Basu, S. (2001) *Weighting socio-economic indicators of human development (a latent variable approach)*. New Delhi: National Institute of Public Finance and Policy.

Njong, M. A. (2007) *Multidimensional Spatial Poverty Comparisons in Cameroon*, Final Report submitted to AERC Research Workshop, Nairobi-Kenya, December. and Wealth, 46, 33–58.

Nolan, B. and Whelan, C.T. (1991) *Resources, Deprivation and the Measurement of Poverty*, The Economic and Social Research Institute; Working Paper No 21.

Lenoir R. (1974), *Les exclus. Un Français sur dix, Le Seuil* or see Saith R., *Social exclusion: the concept and application to developing countries*, Queen Elizabeth House Working Paper Series N°72;

PNUD (1997), *Human Development Report*; New York.

PNUD (2004) *Human Development Report 2004 - Cultural Liberty in Today's Diverse, United Nations Development Programme*, New York.

Ravaillion, M. (1996) *Issues in measuring and modelling poverty*, The Economic Journal, Vol. 106, No. 438, pp. 1328-1343.

Rowena, J. Smith, P. and Goddard, M. (2004) *Measuring Performance: An Examination of Composite Performance Indicators*, CHE Technical Paper Series, 29.

Rowntree, B. S. (1901) *Poverty. A Study of Town Life*; Mc Millan and Co. London. 2nd edition.

Saporta, G. (1983), Multidimensional data analysis and quantification of categorical variables, en *New Trends in Data Analysis and Applications*, J. Janssen, J.F. Marcotorchino, J.M. Proth Eds., Elsevier Science Publishers B.V., North-Holland.

Sen, A. (1993) *Capability and Well-being*, in M.C Nussbaum and A.K Sen (Eds), *The Quality of Life*, Oxford Claredon Press.

Townsend, P. (1979) *Poverty in the United Kingdom, a Survey of Household Resources and Standards of Living*, (London: Penguin Books and Allen Lane).

Townsend, P. (1993) *The International analysis of Poverty*, Harvester Wheatsheat, London, UK

Tsui, K. Y., (2002) *Multidimensional Poverty Indices*, Social Choice and Welfare, Vol 19; pp76-90.

World Bank (2001) *World Development Report 2000/01: Attacking Poverty*, World Bank, Washington D.C.

Young, F.W. (1981), *Quantitative Analysis of Qualitative Data*, Psychometrika, 46, 357-388.

Young, F.W., Takane, Y. y de Leeuw, J. (1978), *The Principal Components of Mixed Measurement Level Multivariate Data: An Alternating Least Squares Method with Optimal Scaling Features*, Psychometrika, 43, 279-281.

Apéndice A

Diagrama de flujo para la construcción del indicador en una dimensión usando los procedimientos PRINQUAL y FACTOR de SAS.

