

CONSTRUCCION DE INDICADORES SOCIALES A PARTIR DE LOS DATOS DE UN CENSO DE POBLACION Y VIVIENDA

**Por Verónica Ibarra G.
Mat. John Ramírez F.**

Lo que se quiere lograr a lo largo de la investigación es construir indicadores sociales, que son estadísticos basados en observaciones que reflejan un aspecto de la vida social que es de interés para la comunidad, particularizando en indicadores de pobreza.

La construcción del indicador se realizará mediante el uso particular de indicadores sintéticos, que es una función matemática de indicadores parciales que reúne un conjunto de condiciones que se juzgan necesarios para concretar una medición expresiva del objetivo buscado.

Las condiciones de pobreza pueden ser identificables en los datos estadísticos de un censo de población y vivienda. Por la información en diversos factores que inciden en la pobreza y por ser parte de las

variables censales es que se decide trabajar con indicadores sintéticos, en los cuales podemos hacer una combinación de estas variables mediante los indicadores parciales que las involucren.

Existen diversas fuentes de datos que pueden utilizarse para evaluar la pobreza, entre ellas se encuentran: las encuestas de hogares, que brindan información para conocer como ha ido evolucionando la pobreza a través del tiempo; además de analizar el impacto productivo de las políticas económicas y sociales utilizadas; así también, las encuestas de ingresos y gastos, encuesta de condiciones de vida y censo de población y vivienda.

El Censo de Población y Vivienda tiene como objetivo dotar de información sobre la

magnitud, estructura, crecimiento, distribución de la población y de sus características demográficas, económicas y sociales, así como de las condiciones generales en que habita la población, que sirva de base para la formulación de políticas de población y para la elaboración de planes de desarrollo y de programas y proyectos a cargo de los organismos públicos y privados.

Los indicadores de pobreza y de bienestar son estadísticos que utilizan datos de encuestas y censos que ayudan a analizar la evolución de la pobreza, a la vez que, permiten hacer comparaciones entre diferentes grupos sociales.

Hay que señalar que los conceptos de pobreza y bienestar están relacionados, pero presentan diferentes matices. El concepto de pobreza estaría relacionado con el ingreso y el gasto de las personas, mientras el concepto de bienestar tiene que ver más con las condiciones de vida, que brinda un nivel adecuado

de satisfacción a las personas, y que no dependen tan directamente del ingreso. Así, los servicios públicos como agua potable, luz eléctrica, alcantarillado, líneas telefónicas, entre otros, brindan cierto grado de bienestar a las personas, pero dependen más de políticas de desarrollo que del ingreso de los hogares

Cada situación requiere una evaluación particular para utilizar el método de análisis multivariado más adecuado que permita tener la máxima información disponible del conjunto de datos, y a su vez garantice la validez de su aplicación.

En el caso particular de la investigación, se quiere explorar la adecuación de las técnicas multivariantes al establecimiento de indicadores sintéticos, sin importar la naturaleza de los datos básicos disponibles. Esta exploración no comenzará con un conjunto de datos a su máximo nivel de desagregación, como sería el individuo, el hogar o el receptor de ingresos. Ello es así porque no se conoce acerca

de los individuos. Para solucionar este problema se trabajará con masas de poblaciones lo que permitirá una buena base experimental.

Está comprobado que el uso de ciertas técnicas es satisfactorio, en el sentido de que las validaciones a que son sometidas a cierto contexto son positivas, queda entonces abierto el camino para su confiable aplicación en otros entornos, con otras condiciones y a otros niveles de agregación que los ensayos en las primeras etapas exploratorias.

La naturaleza multidimensional de los fenómenos sociales en general y de las situaciones y condiciones de pobreza en particular, reclama un tratamiento estadístico que reconozca y respete dichas características. Las técnicas multivariantes logran esto en forma adecuada. Dentro de ellas, la de Componentes Principales, es en principio apropiada para abstraer y revelar estructuras multidimensionales latentes con ciertas propiedades

deseadas. Bajo ciertas condiciones dichas construcciones pueden ser directamente empleadas como indicadores sintéticos del fenómeno o de los fenómenos que el conjunto de variables empleadas busca representar. Un uso juicioso de los resultados que pueden extraerse de su aplicación, brinda la oportunidad para establecer indicadores alternativos de amplia cobertura y capacidad sinterizadora.

El método de Análisis de Componentes Principales (ACP), permite la estructuración de un conjunto de datos multivariantes obtenidos de una población, cuya distribución no es necesariamente conocida.

Lo que se propone es hallar una combinación lineal de las variables representativas del fenómeno multidimensional.

Los objetivos del análisis por el método de componentes principales son:

- ❖ Generar nuevas variables que puedan expresar la

información contenida en el conjunto original de datos.

❖ Reducir la dimensionalidad del problema que se está estudiando, como paso previo para futuros análisis.

❖ Eliminar, cuando sea posible, algunas de las variables originales si ellas aportan poca información.

Las nuevas variables originadas se conocen como *componentes principales*.

Una vez teniendo establecida la técnica que se empleará para la obtención del indicador sintético de pobreza, se procede a la aplicación de la técnica multivariante del Análisis de Componentes Principales

La base de datos usada para el desarrollo del trabajo fue el Resumen de las Principales Variables y Categorías que obtuvo el Instituto de Estadísticas y Censos de Ecuador (INEC), para el V Censo de Población y IV Censo de Vivienda realizado en el año de 1.990.

Se consideraron las 21 provincias existentes hasta el momento censal y se incluye las zonas no delimitadas.

La información de la base de datos obtenida ha sido sintetizada por medio de 15 variables. Éstas son:

X₁ Proporción de viviendas por provincias.

X₂ Indicador de tipo de vivienda.

X₃ Indicador de abastecimiento de agua.

X₄ Indicador del sistema de eliminación de aguas servidas.

X₅ Proporción de las viviendas que disponen de servicio eléctrico del total de viviendas provinciales.

X₆ Proporción de las viviendas que disponen de servicio telefónico del total de viviendas provinciales.

X₇ Indicador del sistema de eliminación de basura.

X₈ Indicador de disponibilidad de servicio higiénico.

X₉ Indicador de disponibilidad de ducha.

X₁₀ Indicador de tenencia de vivienda.

X₁₁ Proporción de la población por provincias

X₁₂ Proporción de la población rural por provincias.

X₁₃ Indicador nivel de instrucción.

X₁₄ Proporción de la población analfabeta de la población total provincial.

X₁₅ Proporción de la población económicamente activa de la población total provincial.

Entre los principales resultados obtenidos tenemos, los valores propios y sus porcentajes de varianza total explicada, lo cual ayudará a determinar cuantos componentes principales son importantes para el análisis de los datos y que sean representativos de las variables originales.

Tabla I
Tabla de Valores Propios obtenidos de la Matriz R y porcentaje de la varianza total explicada

| | Valor Propio | Proporción de la Varianza Total Explicada | |
|----|--------------|---|------------|
| | | Absoluta% | Acumulada% |
| 1 | 9,067 | 60,450% | 60,450% |
| 2 | 2,197 | 14,647% | 75,097% |
| 3 | 1,113 | 7,420% | 82,518% |
| 4 | 1,032 | 6,880% | 89,398% |
| 5 | 0,628 | 4,187% | 93,585% |
| 6 | 0,29 | 1,933% | 95,518% |
| 7 | 0,242 | 1,613% | 97,132% |
| 8 | 0,202 | 1,347% | 98,478% |
| 9 | 0,096 | 0,640% | 99,118% |
| 10 | 0,058 | 0,387% | 99,505% |
| 11 | 0,038 | 0,253% | 99,758% |
| 12 | 0,019 | 0,127% | 99,885% |
| 13 | 0,012 | 0,080% | 99,965% |
| 14 | 0,005 | 0,033% | 99,998% |
| 15 | 2,31665E-04 | 0,002% | 100,000% |

Si se gráfica los valores propios notamos que se puede observar que se produce el “codo” alrededor del cuarto componente, y a su vez estos valores son mayores que 1,

con lo cual se selecciona cuatro componentes principales.

Este número de componentes (q=4), recoge el siguiente

porcentaje de cantidad de información:

$$= \frac{13,409}{15} = 89,398\%$$

*Porcentaje de Varianza
Explicada por los q primeros*

$$\text{componente} = \frac{\sum_{j=1}^q \lambda(j)}{\sum_{j=1}^p \lambda(j)} * 100$$

Los vectores propios asociados a los valores propios seleccionados son:

Tabla II
Tabla de valores propios con sus respectivos vectores propios, basados en el número de factores

| | Valor Propio | | | |
|-----|------------------|--------|--------|--------|
| | 9,067 | 2,197 | 1,113 | 1,032 |
| | Vectores Propios | | | |
| | 1 | 2 | 3 | 4 |
| X1 | 0,230 | 0,38 | 0,103 | -0,393 |
| X2 | 0,045 | -0,312 | 0,666 | -0,294 |
| X3 | 0,28 | -0,163 | 0,182 | 0,319 |
| X4 | 0,319 | 0,071 | 0,05 | 0,068 |
| X5 | 0,307 | -0,038 | 0,143 | 0,203 |
| X6 | 0,273 | 0,066 | 0,355 | -0,128 |
| X7 | 0,288 | -0,162 | -0,005 | 0,231 |
| X8 | 0,298 | 0,07 | -0,125 | 0,04 |
| X9 | 0,324 | -0,039 | -0,107 | 0,094 |
| X10 | 0,187 | 0,416 | -0,107 | 0,375 |
| X11 | 0,226 | 0,399 | 0,061 | -0,387 |
| X12 | -0,307 | 0,072 | 0,177 | 0,049 |
| X13 | 0,309 | -0,176 | -0,11 | -0,102 |
| X14 | -0,159 | 0,25 | 0,522 | 0,477 |
| X15 | 0,139 | -0,507 | -0,076 | -0,033 |

Realizando la respectiva combinación lineal, obtenemos que las nuevas variables-componentes principales- son:

Tabla III

Tabla de componentes principales o factores

| | Comp. I | Comp. II | Comp. III | Comp. IV |
|-----------|----------------|-----------------|------------------|-----------------|
| 1 | 0,145 | -1,397 | 1,696 | -1,492 |
| 2 | -2,275 | -0,146 | 1,904 | 0,015 |
| 3 | -2,305 | -0,496 | 1,672 | -0,362 |
| 4 | 1,812 | -1,068 | 0,778 | 0,937 |
| 5 | -2,101 | 0,798 | 0,408 | 0,878 |
| 6 | -0,689 | 1,318 | 0,393 | 2,134 |
| 7 | 3,037 | -0,273 | -1,760 | 0,699 |
| 8 | -0,984 | 1,241 | -0,261 | 0,168 |
| 9 | 5,742 | 2,997 | -0,436 | -1,774 |
| 10 | 1,478 | -0,071 | 0,917 | 1,599 |
| 11 | -0,956 | 0,119 | -0,822 | -0,453 |
| 12 | -0,438 | 1,664 | -0,755 | 0,690 |
| 13 | 0,731 | 2,198 | -0,084 | 0,045 |
| 14 | -2,865 | -0,836 | -0,314 | -1,037 |
| 15 | -3,468 | -0,134 | -0,979 | -0,599 |
| 16 | -0,125 | -0,366 | -1,317 | 1,238 |
| 17 | 6,932 | 0,557 | 0,813 | -1,056 |
| 18 | 1,544 | -0,877 | 1,075 | 0,464 |
| 19 | -2,885 | -0,708 | -1,076 | -0,709 |
| 20 | 4,635 | -4,402 | -1,039 | 0,196 |
| 21 | -3,459 | -0,498 | -0,914 | -1,031 |
| 22 | -3,506 | 0,383 | 0,101 | -0,549 |

Ya se estableció que el método para construir el indicador social va a ser indicadores sintéticos, para construirlo se aplicará el criterio propuesto por **KAMINSKY**, el cual propone escoger la componente principal de mayor explicación como el

indicador sintético. Es decir, que la Componente Principal I será el indicador sintético de pobreza y las variables constituyen los indicadores parciales.

De esta manera tenemos que las provincias se ubican de la siguiente manera:

Tabla IV
Indicador sintético de pobreza

| Provincias | Indicador de Pobreza |
|----------------------|----------------------|
| Pichincha | 6,93200 |
| Guayas | 5,74200 |
| Galapagos | 4,63500 |
| El Oro | 3,03700 |
| Carchi | 1,81200 |
| Tungurahua | 1,54400 |
| Imbabura | 1,47800 |
| Manabi | 0,73100 |
| Azuay | 0,14500 |
| Pastaza | -0,12500 |
| Los Rios | -0,43800 |
| Chimborazo | -0,68900 |
| Loja | -0,95600 |
| Esmeraldas | -0,98400 |
| Cotopaxi | -2,10100 |
| Bolivar | -2,27500 |
| Cañar | -2,30500 |
| Morona Santiago | -2,86500 |
| Zamora Chinchipe | -2,88500 |
| Sucumbios | -3,45900 |
| Napo | -3,46800 |
| Zonas No Delimitadas | -3,50600 |

Tomado en cuenta la escala, podemos concluir que los que presentan un mayor valor son las provincias con un menor índice de pobreza y viceversa.

Conclusiones:

1. El indicador de pobreza, permite colocar a cada provincia en una posición que dependerá de las otras para interpretarse como “buena” o

“mala”, que ayudará a identificar donde existe retraso en la dotación de servicios, y que evita el desarrollo de la provincia. Este indicador debe ser netamente objetivo para la comunidad y autoridades interesadas, porque aún cuando en el país existen problemas internos, es importante analizar esta posición con la realidad nacional que se vive.

2. La combinación lineal para expresar el Indicador Sintético de Pobreza está dada por la fórmula Indicador

3. Las provincias que obtuvieron una mejor posición fueron Pichincha, Guayas y Galapagos y las últimas Sucumbios, Napo y las zonas no delimitadas.

Recomendaciones:

1. El análisis Aun cuando ha existido desagregación en los datos, resultaría interesante realizar al nivel de cantones y obtener una escala en la cual se pueda notar la concordancia con los indicadores obtenidos en el ámbito provincial. Así también, dividirla en grandes áreas como son: urbana y rural, o por regiones subjetivo del indicador les corresponde a los sociólogos y entendidos de la materia Aun cuando ha existido desagregación en los datos, resultaría interesante realizar al nivel de cantones y obtener una escala en la cual se pueda notar la concordancia con los indicadores obtenidos en el ámbito provincial. Así también, dividirla en grandes

áreas como son: urbana y rural, o por regiones

2. En el Censo de Población y Vivienda del año 1.990 se han notado falencias de información, que podrían llevar a conclusiones erróneas. Este es el caso por ejemplo de los datos de Tenencia de Vivienda, el cual refleja que la mayoría de los pobladores poseen viviendas propias, pero no hay especificaciones sobre que tipo de vivienda propia, si formal o informal, tomando como formal las viviendas sobre las cuales se posean escrituras, caso contrario sería informal.

$$\begin{aligned} \text{Indicador} = & 0.23 * X_1 + 0.045 * X_2 + 0.28 * X_3 + 0.319 * X_4 + 0.307 * X_5 + 0.273 * X_6 + \\ & 0.288 * X_7 + 0.298 * X_8 + 0.324 * X_9 + 0.187 * X_{10} + 0.226 * X_{11} - 0.307 * X_{12} + \\ & 0.309 * X_{13} - 0.159 * X_{14} + 0.139 * X_{15} \end{aligned}$$

BIBLIOGRAFÍA

1. **ABASCAL, Elena.** Métodos Multivariantes para la Investigación Comercial, Ariel Economía, España, 1.989
2. **AGUINAGA, Consuelo.** Determinación de los Niveles de Pobreza Area Urbana del Ecuador, Instituto Nacional de Estadísticas y Censos (INEC), Ecuador, 1.991.
3. **CAPA, Holger.** Modelo para la Evaluación de la Pobreza - Aplicaciones al Ecuador.
4. **DIAZ, Coral.** Indicadores Sintéticos, España. <http://www.festadistica.fguam.es/>
5. **INSTITUTO NACIONAL DE ESTADÍSTICAS Y CENSOS (INEC).** Compendio de las Necesidades Básicas Insatisfechas de la Población Ecuatoriana, Guayaquil, 1.991
6. **INSTITUTO NACIONAL DE ESTADÍSTICAS Y CENSOS (INEC).** Resultados del V Censo de Población y IV de Vivienda de 1.990, Ecuador.
7. **KAMINSKY, Mario.** Vinculación de la Información del Censo de Población y del Censo de Viviendas en Indicadores Sintéticos Alternativos de Condiciones de Pobreza para la Caracterización de Micro Regiones Geográficas - Análisis Cuantitativo Comparativo Multivariante Intra - País e Inter - países: Brasil, Costa Rica, Chile, México, Lima, 1.979
8. **LEE, Haeduck.** Metodología de la Investigación de la Pobreza y la Distribución del Ingreso, Instituto Nacional de Estadísticas y Censos (INEC), Ecuador, 1.994
9. **MEDINA, Fernando.** La evaluación del Bienestar: Alternativas de medición para los Estados de la República mexicana, Estadísticas, México, 1.995, 81 p. FMedina@cap.inegi.gob.mx
10. **PLA, Laura.** Análisis multivariado: Método de componentes principales, Secretaría General de la Organización de los Estados Americanos, Estados Unidos, 1986