

## Construcción de un índice de pobreza a partir del equipamiento de los hogares en Costa Rica

### Construction of a poverty index based on household equipment in Costa Rica

Ángel Jesús Porras Solís<sup>1</sup>  
Universidad Nacional, Costa Rica

**Resumen.** En el presente artículo se aplica una serie de técnicas estadísticas multivariantes para la construcción de un índice de pobreza a partir de los elementos de equipamiento de los hogares en Costa Rica utilizando como fuente de información los microdatos de la Encuesta Nacional de Hogares 2016 (ENAHO 2016). Se calcula el índice y se cuantifica la línea (umbral) que permite separar los hogares pobres y no pobres, para luego estudiar sus repercusiones geográficas (zona y región) y su relación con los otros indicadores de pobreza. Por último, se especifican algunas propiedades que cumple el índice como indicador de medida de pobreza.

**Palabras clave.** Pobreza, equipamiento de los hogares, análisis multivariante, indicadores socioeconómicos.

**Abstract.** In this paper several techniques are applied for the construction of a poverty index based on household equipment in Costa Rica using microdata of the National Household Survey 2016 (ENAHO, 2016). The index is calculated and the line (threshold) that separates poor and non-poor households is quantified, in order to then study its geographical repercussions (zone and region) and its relationship with the other poverty indicators. Finally, are specified some properties that the index meets as an indicator of poverty measurement.

**Keywords.** Ecosystems, coastal zone, forests, mangroves, wetlands, Costa Rica

---

<sup>1</sup>Ángel Jesús Porras Solís. Master en Técnicas Estadísticas, Universidad Santiago de Compostela, España. Coordinador de la Maestría en Desarrollo Rural, Universidad Nacional de Costa Rica. Correo electrónico: [aporra78@gmail.com](mailto:aporra78@gmail.com)

## Introducción

Como bien indican Céspedes y Jiménez (2006, p.7) la pobreza presenta múltiples facetas que hacen imposible una caracterización única y precisa, tanto conceptualmente, como empíricamente cuando se desea medir su magnitud y evolución. En efecto, no existe una definición única y universal de la pobreza, sin embargo, es frecuente encontrar en la literatura especializada, la pobreza definida como una situación donde los hogares no logran un determinado nivel de vida. Al respecto, la mayoría de los estudios económicos sobre pobreza han centrado su atención en definir ese nivel de vida desde el punto de vista de tres concepciones alternativas (Townsend, 1993): subsistencia, necesidades básicas y privación relativa.

Según el concepto de subsistencia, la pobreza se define como la carencia de bienes y servicios materiales (alimentación, vestido y alojamiento) requeridos para vivir y funcionar como un miembro de la sociedad. Toma como punto de partida un cálculo del costo de una canasta de alimentos mínima que cumpla con algunas normas de nutrición y ciertas pautas de consumo de la población.

La pobreza, según el concepto de necesidades básicas, incorpora la visión de que se debe tomar en cuenta que las necesidades básicas son más que las de carácter biológico y que las familias tienen necesidad de satisfacer otro tipo de necesidades como educación, salud, transporte e infraestructura en general (Céspedes y Jiménez, 2006, 7).

Por su parte, la pobreza según el concepto de privación relativa, considera que el modo en que las personas y las familias perciben su estado de pobreza depende de su posición en la sociedad respecto a los demás. Este concepto presta particular atención a que un individuo puede

ser más o menos pobre según cuanto tengan los demás.

En cuanto a su medición, dependiendo del punto de vista adoptado y de los aspectos que interesen resaltar, se pueden realizar diversos análisis de pobreza. Un primer acercamiento se refiere a la pobreza objetiva y subjetiva. Los estudios de pobreza objetiva utilizan información recogida mediante variables cuya medición proviene de la observación directa, como por ejemplo los ingresos y gastos de los hogares. Por su parte, los análisis de pobreza subjetiva “consideran que, por su esencia, las necesidades básicas tienen un componente subjetivo, por lo que las líneas de pobreza deben obtenerse a partir del juicio que las personas tienen sobre los bienes y servicios que consideran esenciales” (Larrain, 2008, 106).

Por medio del enfoque objetivo se realizan tanto análisis de pobreza absoluta como relativa. La pobreza absoluta se define como “la situación en la cual no están cubiertas las necesidades básicas del individuo, es decir, existe carencia de bienes y servicios básicos, normalmente relacionados con la alimentación, la vivienda y el vestido” (INE, 2007, 3). Dicho enfoque cuenta con métodos de medición tales como el de Necesidades Básicas Insatisfechas (NBI) y línea de pobreza. El método de NBI cumple como objetivo identificar los hogares pobres en función de un conjunto de necesidades socialmente consideradas básicas, de manera que se clasifican como hogares pobres aquellos que no logran satisfacer los umbrales mínimos determinados para cada necesidad; por consiguiente, se busca la identificación y caracterización de los hogares que no satisfacen aquellas necesidades humanas que dependen de condiciones económicas (INEC, 2015, 11). El método de línea de pobreza, compara los ingresos de los hogares con el costo de una canasta de bienes y servicios requeridos para mantener un nivel mínimo de subsistencia

(INEC, 2010, 19). Se considera en condición de pobreza los hogares que no logaran cubrir con sus ingresos, el costo de la canasta básica de bienes y servicios.

La pobreza relativa sitúa el fenómeno de la pobreza en la sociedad objeto de estudio. Desde esta perspectiva se considera que una persona es pobre cuando se encuentra en una situación de clara desventaja, económica y socialmente, respecto al resto de personas de su entorno desigualdad (INE, 2007, 3). De acuerdo a este enfoque “las necesidades surgen de la comparación con el resto de la sociedad, por ello, las líneas de pobreza que emanan de aquí se fijan en relación al ingreso medio de un país” (Larrain 2008, 108). Por ejemplo, en países como España, el umbral de pobreza se define como 60% de la renta mediana, de forma que si el ingreso de un hogar se encuentra por debajo de este nivel estará por debajo del umbral de pobreza.

En Costa Rica, la medición anual de la pobreza se realiza desde 1987 aplicando el método de línea pobreza (INEC, 2010, 20). No obstante, a partir del 2015 se incorpora otra medición anual conocida como Índice de Pobreza Multidimensional, la cual identifica las necesidades que presentan los hogares en cinco dimensiones consideradas prioritarias para evaluar la calidad de vida: vivienda y uso de Internet, salud, educación, trabajo y protección social; cada dimensión se compone de cuatro indicadores o carencias que sufren las personas (INEC, 2015).

No existe por el momento un consenso acerca de cuál es el mejor índice entre todos los propuestos hasta el presente. Por tal razón, en los últimos años se viene recalando la necesidad de proporcionar otras medidas de pobreza, que consideren además de los indicadores monetarios, variables que reflejen directamente las privaciones que sufren los hogares.

La creciente necesidad de proporcionar medidas no monetarias de privación tiene múltiples causas: por un lado, la pobreza monetaria muestra solo una parte del fenómeno y presupone que hogares que tienen los mismos ingresos gozan de niveles de vida similares. Por otra parte, los individuos poseen otro tipo de recursos, no reflejados en las mediciones de pobreza monetaria, que pueden utilizar para no caer en la pobreza y llevar un nivel de vida aceptable (INE, 2007, 27).

Considerando lo anterior, en el presente artículo se aplica una metodología para la determinación de los niveles de pobreza, así como la identificación de los hogares en dicha situación, desde el punto de vista de la dotación de bienes y servicios de los hogares. Esto permite conocer el fenómeno de la pobreza desde una perspectiva multidimensional, a través de un indicador no monetario. Bajo esta perspectiva, se elabora un índice de pobreza mediante la aplicación de diversas técnicas estadísticas multivariantes. De igual forma, se procede a la aplicación empírica del índice utilizando datos provenientes de la Encuesta Nacional de Hogares (ENAH) correspondiente al 2016.

La organización del artículo contempla dos partes. La primera describe la metodología seguida para la construcción del índice, enfatizando en las distintas técnicas estadísticas multivariantes utilizadas. En primera instancia, para obtener una ordenación y clasificación inicial de los hogares y reducir la dimensionalidad del conjunto de variables, se obtiene un indicador sintético de distancia. En función de lo anterior, se procede a realizar un análisis discriminante, con el fin de utilizar como índice de pobreza las puntuaciones de la primera función canónica discriminante. Para determinar el punto de corte, que permita diferenciar los hogares pobres de los hogares no pobres (línea o umbral de pobreza), se identifican un grupo de bienes y

servicios de primera necesidad o básicos en los hogares (dotación mínima) mediante un análisis de clúster.

La segunda parte está dedicada propiamente a los resultados y discusión. Primero se calcula el índice de pobreza y se cuantifica la línea de pobreza, luego se estudian sus repercusiones regionales (según zona y región) y su relación con los otros indicadores de pobreza calculados por el Instituto Nacional de Estadísticas y Censos de Costa Rica (INEC) en la misma ENAHO 2016. Por último, se indican algunas propiedades que cumple el indicador de pobreza propuesto.

Es importante mencionar, que el procedimiento para la selección de variables, método de reducción de dimensionalidad, cálculo del índice y determinación de la línea de pobreza, se fundamenta en su totalidad en la metodología propuesta por García, Sánchez y Vaamonde (2007) en "A pobreza en Galicia. Medición a través do equipamento dos fogares" y que el manejo de los datos, rutinas de cálculo, algoritmos, técnicas multivariantes y demás procedimientos estadísticos se realizan con el programa estadístico R.

## **Materiales y métodos**

Para construir el indicador de pobreza basado en los elementos de equipamiento de los hogares, se inicia con una selección de variables correspondientes al módulo de vivienda y servicios de la ENAHO 2016. De acuerdo a García et al. (2007, 104-105) la selección de las variables se condiciona a la disponibilidad de información de las propias encuestas de hogares y por la operatividad del modelo propuesto para el cálculo del índice, en cuanto a seleccionar variables que indiquen posesión o no, de un determinado bien o servicio, es decir, variables dicotómicas. En razón de lo anterior, se seleccionan un total de 18 variables: suministro de agua potable, tenencia de servicio sanitario, tenencia

de baño, suministro de electricidad, teléfono celular, teléfono residencial, refrigeradora, sistema de agua caliente, tanque para almacenar agua, computadora portátil, computadora de escritorio, tableta, radio o equipo de sonido, tv de plasma/lcd/led, tv convencional, televisión de paga por cable, televisión de paga por satélite y conexión a Internet. Es necesario aclarar que se toman en cuenta los elementos de equipamiento y servicios que registra la ENAHO 2016; no está considerado la tenencia de elementos como lavadora, horno de microondas o reproductores de video, que si se registraba antes del 2010 en las Encuestas de Hogares y Propósitos Múltiples. La inclusión y medición de cada uno de los elementos de equipamiento que se incluyen en la ENAHO, de acuerdo al INEC (2016), se justifica cuando lo que interesa registrar es la disponibilidad que se tiene del artefacto o servicio y proveer información para determinar grados de accesibilidad a bienes y servicios, así como ser un insumo en estudios de estratificación y exclusión social, entre otros motivos.

Cada variable se recodifica como 1: tenencia y 0: no tenencia. Para todas las variables de equipamiento y la variable tenencia de baño, puesto que las respuestas son sí o no, la recodificación es automática (sí: 1, no: 0). La variable suministro de electricidad toma el valor de 1 si la electricidad es suministrada por alguna institución, organización o empresa del sector y 0 si no es eléctrico o no se tiene. La fuente del agua de consumo (suministro de agua potable) toma el valor de 1 si el agua es suministrada por alguna institución, organización o empresa del sector y 0 si el agua proviene de río, quebrada o naciente, de la lluvia o de otra fuente. La tenencia de servicio sanitario, toma el valor de 1 si el sistema de disposición de excretas es alcantarillado, cloaca, tanque séptico, pozo negro o letrina y 0 si no se tiene, o si el sistema es diferente a los mencionados. Las dos categorías de tenencia de televisor (convencional o led, lcd,

plasma, etc.) se agrupó en una sola variable, al igual que el servicio de televisión de paga (cable o satélite). Es importante advertir que la propuesta de García et al. (2007) para el cálculo del índice de pobreza, no considera la cantidad de ciertos artefactos que se disponen en el hogar (por ejemplo, la cantidad de teléfonos celulares), lo cual puede ser considerado como una limitación del modelo, pero también puede ser tomado como una importante variante a incorporar en futuros estudios similares.

Como unidad de análisis se utiliza el hogar, que se entiende como “personas con vínculos familiares o sin ellos, que residen habitualmente en una vivienda individual, que participan de la formación, utilización, o ambas, de un mismo presupuesto, que llevan una vida en común, y elaboran y consumen en común sus alimentos” (INEC, 2016). La muestra de hogares que se utiliza es de 11.335, que corresponde a todos los hogares que efectivamente encuestó el INEC para la ENAHO 2016 (ver tabla 1).

Para obtener una ordenación y clasificación inicial de los hogares y reducir la dimensionalidad del conjunto de variables, se aplica la medida de distancia DPM propuesta por García et al. (2007). Las medidas de distancia constituyen un enfoque muy habitual para el diseño de indicadores sintéticos socioeconómicos relacionados con la pobreza y la desigualdad. Una de las medidas de distancia más conocidas y aplicadas es la  $DP_2$ , diseñada por el economista español Jesús Pena Trapero (1977) para medir el bienestar social a través de un indicador sintético que agrega la información contenida en un conjunto de indicadores sociales y que está diseñada para realizar comparaciones entre diferentes unidades en un momento dado de tiempo. Se toma como referencia una unidad teórica que puede alcanzar los valores máximos o mínimos (valor norma o referencia) de las variables objeto de estudio.

Las unidades pueden ser países, regiones, hogares, etc. El indicador  $DP_2$  devuelve las

Tabla 1

*Muestra de hogares a utilizar para el análisis empírico, por zona y región de planificación. ENAHO 2016*

Zona y región	Total
Costa Rica	11.335
Zona	
Urbano	7.873
Rural	3.462
Región	
Central	7.868
Chorotega	1.152
Pacífico Central	912
Brunca	1.152
Huetar Atlántica	1.248
Huetar Norte	1.008

*Fuente.* Instituto Nacional de Estadísticas y Censos (INEC). ENAHO 2016



distancias de cada unidad respecto a esa unidad teórica de referencia. Se define  $DP_2$  de la siguiente forma (Zarzosa, 2009):

$$DP_2 = \sum_{i=1}^n \frac{d_i}{\sigma_i} (1 - R_{i, i-1, \dots, 2, 1}^2)$$

donde

$d_i = d_i(r^*) = |x_{ri} - x_{*i}|$ ; es decir, la diferencia en valor absoluto, entre el valor de la variable  $i$  ( $i = 1, \dots, n$ ) y el valor norma o referencia de dicha variable y además tipificado por la inversa de la desviación típica  $\sigma_i$  correspondiente a los valores mostrados por la variable  $i$ .

$R_{i, i-1, \dots, 2, 1}^2$  es el coeficiente de determinación de la regresión lineal de cada variable con respecto a las anteriores.

De acuerdo a García et al. (2007, 116), el hecho de trabajar con variables dicotómicas, donde 0 denota la ausencia de un determinado elemento de equipamiento del hogar y 1 la presencia del mismo, sugiere modificar algunos de los procedimientos de cálculo de la  $DP_2$ . Básicamente se trata de redefinir los valores norma y utilizar regresión logística para obtener los  $R^2$ . Como base de referencia para cada variable se sugiere el valor 0, correspondiente a la situación teórica menos deseada, es decir, la ausencia de un determinado elemento de equipamiento o servicio en el hogar. Considerando lo anterior, se propone una medida de distancia DPM:

$$DP_M = \sum_{i=1}^n d_i (1 - R_{i, i-1, \dots, 2, 1}^2)$$

donde

$d_i = d_i(r^*) = |x_{ri} - x_{*i}|$ ; es decir, la diferencia en valor absoluto, entre el valor de la variable  $i$  ( $i = 1, \dots, n$ ) y el valor norma, que es igual a 0.

$R_{i, i-1, \dots, 2, 1}^2$  es el coeficiente de determinación

de la regresión logística de cada variable con respecto a las anteriores.

De esta forma, la clasificación de los hogares se realiza en función del valor de la DPM, donde serán clasificados con valores bajos aquellos hogares que carecen de prácticamente todo el equipamiento, mientras que en el extremo, aparecen los hogares con mayor equipamiento.

En función de los resultados obtenidos con DPM, se procede a un análisis discriminante, el cual procura resumir en un índice toda la información contenida en diversas variables independientes. Para lograr dicha finalidad se busca estimar el peso de las variables, que permita lograr la mejor separación posible entre grupos, así como obtener una función canónica discriminante, utilizable como índice de pobreza IPM.

El análisis discriminante como técnica estadística de clasificación se utiliza para describir las características que distinguen a los miembros de un grupo (sujetos u objetos) y proporcionar una regla discriminante que permita asignar un nuevo individuo u objeto de manera óptima a una clase o grupo previamente identificado, donde esta regla discriminante se obtiene utilizando la información de un número determinado de variables de una muestra de individuos correspondiente a cada uno de las clases o grupos identificados (Hernández, 2013, 163).

Con respecto a la muestra de hogares para conformar estos grupos y aplicar el correspondiente análisis discriminante de acuerdo al indicador propuesto, García et al. (2007, 118-119) señalan que se debe identificar dos grupos extremos, más un grupo intermedio. Específicamente, se sugiere tomar 10% de los hogares con menor valor DPM y 10% con mayor el valor, así como los situados entre los percentiles 45 y 55 del DPM. Así, se dispondrán de una muestra de hogares de similar tamaño y que reflejan tres situaciones posibles.

Para cada uno de los hogares se calcula un puntaje discriminante haciendo uso de los coeficientes de la función, los cuales se multiplican por el correspondiente valor de la variable y luego se suman los productos, obteniendo el índice de pobreza IPM.

$$IPM_k = \alpha_1 x_{1k} + \alpha_2 x_{2k} + \dots + \alpha_p x_{pk}$$

donde

$IPM_k$  = puntuación IPM de la función discriminante para el hogar  $k$ ,

$\alpha_0$  = ponderación discriminante para la variable independiente  $p$ ,

$x_{pk}$  = valor de variable independiente  $p$  para el hogar  $k$ .

Una vez determinado el puntaje, se utiliza la regla de Bayes para estimar la probabilidad de que un hogar en particular con un puntaje discriminante pertenezca a un grupo en particular. Los resultados de la clasificación, se resumen en una matriz de confusión, que da una idea de la tasa de clasificaciones incorrectas.

Para valorar la significación estadística de las funciones canónicas discriminantes, se utiliza la Lambda de Wilks, cuyos valores varían entre 0 y 1. Así, valores próximos a 0 indican mucha discriminación, es decir los grupos están muy separados, mientras que valores cercanos a 1 representan escasa discriminación o poca diferencia entre los grupos. Si la función discriminante es estadísticamente significativa y su capacidad para clasificar es aceptable, se pueden realizar interpretaciones de sus resultados, donde se puede determinar la importancia relativa de cada una de las variables independientes en su capacidad para discriminar entre los grupos.

También en el modelo discriminante se evalúa la posición de los centroides de grupo. Los centroides son los valores medios para cada uno de los grupos en cada una de las funciones discriminantes y por lo tanto, se utiliza también para conocer cuál función discrimina mejor (Díaz Díaz, González y Henao, 2013, 193).

Para llevar a cabo el análisis discriminante, las variables se incluyen en el modelo simultáneamente, ya que el objetivo es determinar las capacidades discriminantes del conjunto completo de variables, sin considerar la influencia de ninguna individualmente. A su vez se utilizan las aplicaciones de R, *Canonical Discriminant Analysis* (candisc) por Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M. y Hornik, K. (2013) y *Linear Discriminant Analysis* (lda) por Venables y Ripley (2002).

Una vez que se tiene la función canónica discriminante para obtener las puntuaciones del IPM en cada uno de los hogares, lo siguiente es determinar un punto de corte en las puntuaciones, que permita diferenciar los hogares pobres de los hogares no pobres. Esto implica en primera instancia, identificar un grupo de bienes y servicios de primera necesidad o básicos en los hogares, a lo que García et al. (2007, 138) denominan dotación mínima.

Efectivamente, el análisis de clúster permite identificar esta dotación mínima de bienes y servicios que requieren los hogares. Es una técnica cuya idea básica es agrupar un conjunto de variables en un número dado de clúster. Dicho agrupamiento se basa en la idea de similitud entre las variables, que es una medida de correspondencia entre objetos que van a ser agrupados. En el caso particular de las variables dicotómicas, cuyas características se miden en términos no métricos, se suele utilizar coeficientes de similitud calculados a partir de las frecuencias de tenencia del bien o servicio en el hogar.

A su vez, el análisis de clúster requiere un algoritmo de clasificación que permita agrupar a las variables similares. En ese sentido, se sigue un algoritmo de clasificación jerárquico de tipo aglomerativo, que agrupa las unidades de análisis en cada paso hasta llegar a un conjunto que engloba a la totalidad de las variables. Finalmente se obtiene una representación gráfica conocida como dendrograma, el cual tiene forma de árbol y resume el proceso de agrupación. Los objetos similares se conectan mediante enlaces cuya posición en el diagrama está determinada por el nivel de similitud entre los objetos. Para realizar el análisis de clúster jerárquico se utiliza la aplicación de R, *Hierarchical Clustering* (hclust)

por R Core Team (2013). A partir de esta dotación mínima, se cuantifica el valor del IPM que separa los hogares pobres de los hogares no pobres a partir de los valores correspondientes de los coeficientes no estandarizados de la primera función canónica discriminante.

### Resultados y discusión

Primero, se realiza un breve análisis de frecuencias relativas de las variables, con la intención de contar con una idea preliminar de los datos y su magnitud a utilizar a lo largo de la aplicación empírica. Los mismos hacen referencia al porcentaje de tenencia de un determinado bien o servicio objeto del presente trabajo (ver tabla 2).

Tabla 2  
*Frecuencias relativas de las variables*

Variable	Porcentaje
Suministro de agua potable	95,34
Tenencia de servicio sanitario	99,72
Tenencia de baño	99,33
Suministro de electricidad	99,48
Teléfono celular	95,73
Teléfono residencial	39,80
Refrigeradora	95,29
Sistema para agua caliente para toda la casa	6,42
Tanque para almacenar agua	12,13
Computadora portátil	38,01
Computadora de escritorio	19,54
Tableta	24,45
Equipo de sonido o radio	65,18
Televisor	97,13
Televisión de paga	66,41
Conexión a Internet	64,92

*Fuente.* Instituto Nacional de Estadísticas y Censos (INEC). ENAHO 2016



Entre los principales resultados se puede destacar que casi todos los hogares disponen de los servicios básicos como agua potable, electricidad, baño y sanitario. De igual forma, se observa que la mayoría cuenta con refrigeradora y televisión. Un porcentaje muy bajo (6,42%) cuenta con sistema de agua caliente para toda la casa y un 12,13% disponen de tanque de almacenamiento de agua. También se observa que 95,73% de los hogares cuentan con teléfono celular y que el servicio de teléfono residencial apenas representa cerca de 40%. La tenencia de computadoras portátil registra 38,1%, casi el doble que la tenencia de computadora de escritorio. El porcentaje de hogares con acceso a Internet alcanza 64,92%. Por último, un porcentaje muy bajo de los hogares cuentan con televisión por satélite. No obstante, un poco más de la mitad de los hogares cuentan con televisión de pago por cable.

Como primer paso para obtener una ordenación de los hogares, se realiza el cálculo de la medida de distancia DPM. Así, la secuencia de la DPM constituye la base para aplicar un análisis

discriminante, que se centra en la obtención de una función canónica discriminante, que se pueda utilizar como índice de pobreza IPM, mediante la conformación de tres grupos diferenciados de hogares: Grupo 0: los más pobres, Grupo 1: grupo intermedio y Grupo 2: los menos pobres. Una vez obtenidos los percentiles, los tres grupos quedan conformados por 1.171, 1.195 y 1.134 hogares, respectivamente. En efecto, con la anterior muestra de hogares se obtiene la función canónica discriminante, la cual luego se aplica a los 11.335 hogares que conforma la ENAHO 2016, esto con el propósito de tener un valor de IPM para cada hogar.

Luego de aplicar el análisis discriminante, se procede a valorar la calidad del procedimiento de clasificación. En la tabla 3 se puede constatar que el 99,9% de los casos son agrupados correctamente. Los pocos hogares clasificados incorrectamente del grupo 0 son incluidos en el grupo 1 y también un hogar del grupo 1 es incluido en el grupo 2, el porcentaje de hogares clasificados correctamente en el grupo 2 es de 100%.

Tabla 3  
*Resultados de la clasificación*

		Grupo de pertinencia pronosticado			Total
		0	1	2	
Recuento	0	1.162	9	0	1.171
	1	0	1.194	1	1.195
	2	0	0	1.134	1.134
Porcentaje	0	99,23	0,77	0	100
	1	0,25	99,91	0,09	100
	2	0	0	100	100

*Fuente.* Elaboración propia.

En la figura 1 se pueden ver las proyecciones de los hogares sobre las dos primeras funciones canónicas discriminantes, donde se distinguen claramente la posición de cada uno de los tres grupos en el plano, a la izquierda el grupo 0 (más pobre), a la derecha el grupo 2 (menos pobre) y el centro el grupo 1 (intermedio). En términos generales, los tres grupos aparecen bien separados.

El análisis de la posición relativa de los centroides en los distintos ejes discriminantes, también permite obtener una perspectiva global de la separación entre los grupos (Díaz et al., 2013, 193). El centroide del grupo 0 en el eje definido por la primera función discriminante es -6,73, mientras que el centroide del grupo 2 (7,06) se encuentra en el extremo opuesto. El grupo 1, cuyo centroide es -0,14, se sitúa en una posición intermedia cercana al origen.

En la tabla 4 se observa que la capacidad discriminante de la primera función, con respecto al total acumulado es de 95,03%, mientras que el porcentaje relativo correspondiente a la segunda función es solamente de 4,97%. Esto indica que la primera función es más potente; es decir, aporta la mayor parte de la información para la separación de grupos. De manera similar, los valores del autovalor corroboran que la primera función cumple mejor el objetivo de separar los grupos.

Para valorar la significación estadística de las funciones discriminantes, de acuerdo al valor de Lambda de Wilks, las dos resultan significativas (ver tabla 5). Sin embargo, como se mencionó anteriormente, el autovalor correspondiente a la primera función tiene un peso relativo de 95,03%, de modo que la segunda función, aun siendo estadísticamente significativa, tiene poca importancia práctica a la hora de separar los grupos.

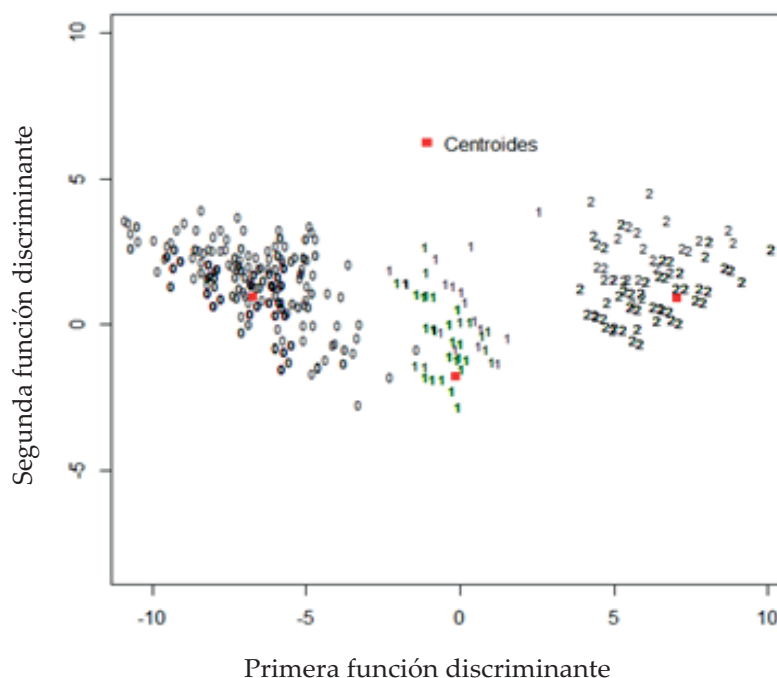


Figura 1. Puntuaciones discriminantes. Fuente: Elaboración propia.

Tabla 4  
*Resumen de las funciones canónicas discriminantes*

Función canónica	Autovalor	Porcentaje de varianza explicada
1	31,43	95,03
2	1,64	4,97

Fuente. Porras Solís, 2018

Tabla 5  
*Lambda de Wilks*

Contraste de las funciones	Lambda de Wilks	p-valor
1 a la 2	0,01	0,000
2	0,37	0,000

Fuente. Porras Solís, 2018

Con el objetivo de determinar la importancia relativa de cada variable para discriminar entre los grupos, se pueden examinar las ponderaciones discriminantes estandarizadas y las cargas de las funciones que resultan del análisis (tabla 6).

Atendiendo en primer lugar al valor absoluto de los coeficientes estandarizados de la primera función, se puede constatar que la variable televisión de paga es la de mayor importancia a la hora de predecir el grupo de pertenencia de los hogares, seguida por las variables teléfono residencial, equipo de sonido o radio y computadora portátil, por su parte las variables de menor importancia son la tenencia de baño, sanitario y televisor.

Tabla 6  
*Coefficientes de las funciones canónicas discriminantes*

Variables	No estandarizados	Estandarizados	Estructura
Suministro de agua potable	1,38	0,35	0,32
Tenencia de servicio sanitario	0,96	0,08	0,11
Tenencia de baño	0,42	0,05	0,15
Suministro de electricidad	-1,10	-0,14	0,17
Teléfono celular	1,18	0,28	0,32
Teléfono residencial	2,29	0,71	0,73
Refrigeradora	1,30	0,36	0,42
Sistema para agua caliente para toda la casa	0,92	0,25	0,46
Tanque para almacenar agua	1,50	0,50	0,42
Computadora portátil	2,13	0,67	0,73
Computadora de escritorio	2,04	0,62	0,68
Tableta	1,27	0,39	0,64
Equipo de sonido o radio	2,02	0,70	0,65
Televisor	-0,22	-0,05	0,30
Televisión de paga	2,33	0,75	0,74
Conexión a internet	2,42	0,59	0,77

Fuente. Porras Solís, 2018

Con respecto a los coeficientes de estructura, los cuales miden la correlación entre las variables y las funciones canónicas discriminantes estandarizadas, se puede apreciar como la variable conexión a internet es la que mayor correlación tiene con la primera función. Las variables tenencia de sanitario y baño, así como suministro de electricidad registran los valores más bajos. En la figura 2 se muestra de una forma gráfica las magnitudes de estas correlaciones, así como la distribución de las puntuaciones de cada grupo obtenidas a partir de la primera función canónica discriminante.

Por lo tanto, el índice de pobreza IPM corresponderá a las puntuaciones obtenidas a partir de la primera función canónica discriminante (primera columna de la tabla 6). Para una mejor comprensión y un manejo óptimo de los puntajes, estos se transforman en valores en escala de 0 a 100, donde 0 corresponde al puntaje más bajo y 100 al puntaje más alto. Para

lograr lo anterior se aplica la fórmula propuesta por Fournier y Garner (1990):

$$X_{NE} = (X_{VE} - Li_{VE}) \left( \frac{Ls_{NE}}{Ls_{VE} - Li_{VE}} \right)$$

donde

$X_{NE}$  = puntaje de la variable con una nueva escala y límite inferior cero

$X_{VE}$  = puntaje de la variable original (vieja escala)

$Ls_{NE}$  = límite superior teórico de la variable con nueva escala y límite inferior cero

$Li_{VE}$  = límite inferior teórico de la variable original (vieja escala)

$Ls_{VE}$  = límite superior teórico de la variable original (vieja escala)

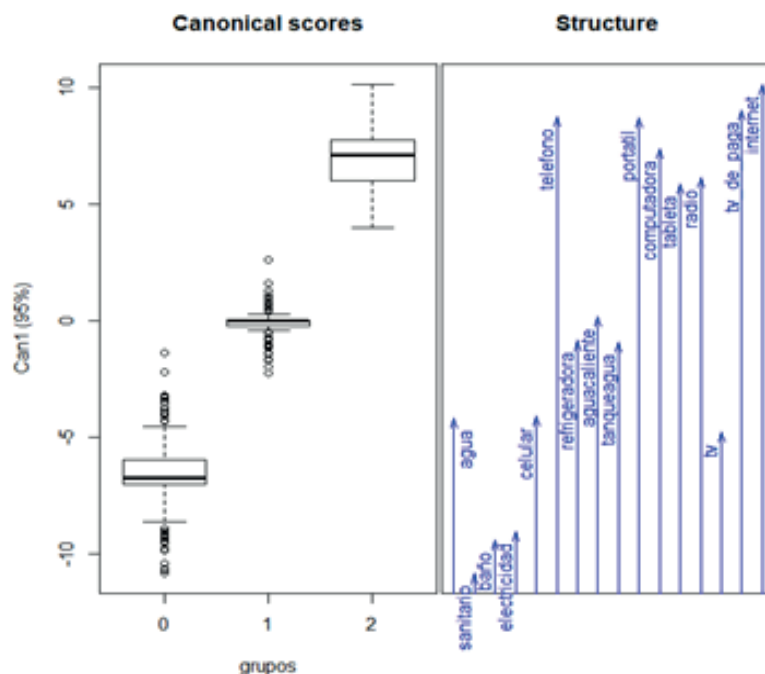


Figura 2. Puntuaciones y coeficientes de estructura de la primera función canónica discriminante. Fuente: Porras Solís, 2018

En la tabla 7, se muestra que el IPM tiene un valor promedio de 51,44 con una desviación típica de 19,70 lo cual se traduce en un coeficiente de variación relativamente alto (38,29%). De igual forma los valores correspondientes a los tres percentiles y su distancia entre sí, evidencian que el IPM sigue un recorrido amplio.

Una vez que se obtiene la función canónica discriminante que será utilizada para obtener las puntuaciones del IPM en cada uno de los hogares, lo siguiente es determinar un punto de corte en las puntuaciones, que permita diferenciar los hogares pobres de los hogares no pobres. Esto implica identificar un grupo

de bienes y servicios básicos en los hogares (dotación mínima) y luego cuantificar el valor de IPM que justamente, separa los hogares pobres de los hogares no pobres. Lo anterior se resuelve con un análisis de clúster.

De acuerdo al análisis clúster, cuyo resultado se resume en el dendrograma de la figura 3, es posible distinguir un primer grupo de bienes y servicios que pueden ser considerados básicos en los hogares (dotación mínima): suministro de agua potable, tenencia de servicio sanitario, tenencia de baño, suministro de electricidad, refrigeradora, televisor y teléfono celular. El resto de los bienes y servicios, agrupados en

Tabla 7  
Medidas estadísticas de IPM

Medida estadística	Valor
Media	51,44
Desviación típica	19,70
Coeficiente de variación	38,29%
Mínimo	0
Percentil 25	38,68
Mediana (Percentil 50)	51,57
Percentil 75	62,52
Máximo	100

Fuente. Porras Solís, 2018

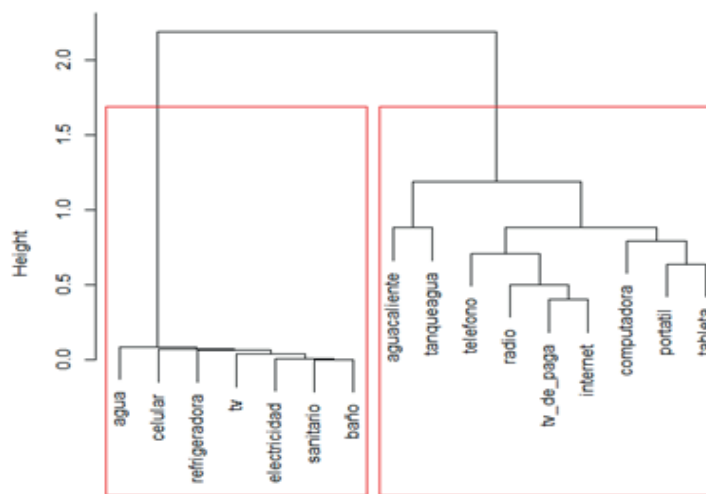


Figura 3. Dendrograma. Análisis de clúster jerárquico. Fuente: Porras Solís, 2018



un segundo clúster, serían aquellos que una vez alcanzada la dotación mínima, contribuyen mediante su tenencia o disponibilidad a incrementar las comodidades de los hogares.

Al tomar como referencia el grupo de variables que conforman la dotación mínima y los correspondientes valores de la primera función canónica discriminante para cada variable contenidos en la primera columna de la tabla 6 y al aplicar la respectiva transformación de escala, se obtiene un valor crítico para clasificar los hogares y se establece que la línea de pobreza se sitúa en el valor  $IPM = 22,05$  que refiere al valor más alto de los bienes y servicios que conforman la dotación mínima. De esta manera, los hogares con un IPM igual o menor a dicho valor, serán clasificados como pobres y los que registren un valor más alto, serán considerados como no pobres.

La línea de pobreza anteriormente establecida, permite concluir que 6,23% de los hogares se encuentran por debajo de la misma. Por tanto, éste sería el porcentaje de hogares clasificados como pobres. Consecuentemente, 93,77% de los

hogares restantes se sitúa por arriba de dicho límite, y resultan clasificados como no pobres.

Dado que uno de los objetivos del trabajo es determinar la repercusión del fenómeno de la pobreza a lo largo del territorio nacional, precisamente el diseño muestral de la ENAHO permite realizar dicho análisis desde dos ámbitos: zona y región (ver tabla 8). Para ello, se toma como referencia la línea de pobreza obtenida anteriormente a nivel nacional: se analiza el porcentaje de hogares considerados como pobres en cada uno de los ámbitos (zona y región). Asimismo, se aplicaron los respectivos factores de expansión para la representatividad de los datos a nivel nacional, lo cual implica expandir la muestra de 11.335 hogares de la muestra a un total poblacional de 1.496.829 hogares. El factor de expansión es un ponderador que se aplica a cada unidad de estudio en la muestra para obtener una estimación poblacional; se interpreta como la cantidad de unidades en la población que representa cada unidad en la muestra, ya sea vivienda, hogar o persona (INEC, 2016, 25).

Tabla 8  
*Cantidad y porcentaje de hogares pobres por zona y región*

Zona y región	Cantidad	Porcentaje
Costa Rica	93.252	6,23
Zona		
Urbano	42.748	3,92
Rural	50.504	12,43
Región		
Central	31.448	3,38
Chorotega	10.809	9,46
Pacífico Central	6.876	7,65
Brunca	9.584	8,40
Huetar Atlántica	19.299	14,54
Huetar Norte	15.236	13,13

*Fuente.* Porras Solís, 2018

Los resultados muestran que la pobreza afecta en mayor medida a los hogares ubicados en las zonas rurales, en particular los ubicados en las regiones Huetar Atlántica y Huetar Norte. La Región Central registra el porcentaje de hogares con menos pobreza, seguida por las regiones Pacífico Central, Brunca y Chorotega, con porcentajes inferiores a 10%. En términos de cantidad de personas, significa un total de 251.737, de quienes 55% se ubican en la zona rural.

En referencia a los valores del IPM por región, en el diagrama de caja de la figura 4, se reflejan tales diferencias entre los valores registrados por la región Central con respecto a las demás,

en especial con las regiones Huetar Atlántica y Huetar Norte.

No obstante, como se puede observar también en la figura 4, en estas regiones se pueden encontrar algunos hogares con altos valores del IPM, los cuales se pueden considerar como atípicos, respecto a la región a la cual pertenecen. Si estos últimos resultados se complementan con un estudio de desigualdad por medio del cálculo del índice de Gini para los valores el IPM de acuerdo a la zona y región, se obtienen prácticamente los mismos resultados, y una mayor desigualdad en la zona rural y en las regiones Huetar Norte y la Huetar Caribe (ver tabla 9).

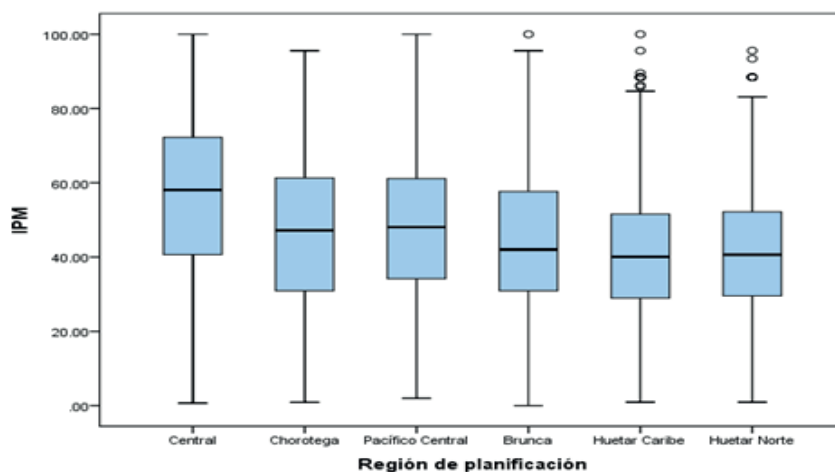


Figura 4. Diagrama de caja del IPM por región. Fuente: Porras Solís (2018).

Tabla 9

Índice de Gini por zona y región

Zona y región	Índice de Gini
Zona	
Urbano	0,20
Rural	0,23
Región	
Central	0,20
Chorotega	0,21
Pacífico Central	0,20
Brunca	0,21
Huetar Atlántica	0,24
Huetar Norte	0,23

Fuente. Porras Solís, 2018

En cuanto a la distribución de los hogares pobres (IPM), respecto a su nivel de ingresos expresados mediante los quintiles de ingresos per cápita, en la tabla 10 se puede observar que 71,1% de los hogares pobres se encuentran en los dos primeros quintiles, lo cual indica, como es de esperar, una alta relación entre los ingresos que perciben los hogares y la tenencia de equipamiento. El INEC divide los hogares en cinco grupos, según su nivel de ingresos. Cada grupo se denomina quintil y conforme aumenta el nivel de ingresos de los hogares, crece el número de quintil en el cual se ubican.

Así el quintil I contiene a los hogares de menores ingresos y el quintil V a los que perciben más.

Si se examina la correspondencia que existe entre los hogares pobres identificados por medio del IPM, con respecto a los hogares en condición de pobreza por insuficiencia de ingresos y pobreza multidimensional (ambos calculados por el INEC), como se muestra en la tabla 11, a nivel nacional 43,2% de los hogares pobres (IPM) se encuentran también en condición de pobreza por insuficiencia de ingresos; mientras que en términos de pobreza multidimensional, significa 53,7% de los hogares. En Costa Rica se mide el IPM a partir del 2015 e identifica las

Tabla 10

*Distribución de los hogares pobres (IPM) por quintil de ingresos per cápita*

Quintil	Ingresos en colones	Porcentaje de hogares
I	94.063 o menos	47,3
II	Más de 94.063 a 165.618	23,8
III	Más de 165.618 a 273.816	16,1
IV	Más de 273.816 a 499.386	9,7
V	Más de 499.386	3,2

Fuente. Porras Solís, 2018

Tabla 11

*Hogares pobres IPM por zona y región. Porcentaje de hogares con pobreza por insuficiencia de ingresos y pobreza multidimensional*

Zona y región	Pobreza por insuficiencia de ingresos	Pobreza multidimensional
Costa Rica	43,2	53,7
Zona		
Urbano	47,0	45,7
Rural	39,9	60,6
Región		
Central	39,9	46,3
Chorotega	36,9	50,0
Pacífico Central	47,4	52,1
Brunca	51,2	44,7
Huetar Atlántica	39,1	57,4
Huetar Norte	52,5	73,6

Fuente. Porras Solís, 2018

necesidades que presentan los hogares en cinco dimensiones, consideradas prioritarias para evaluar la calidad de vida: vivienda y uso de internet, salud, educación, trabajo y protección social, donde cada dimensión se compone de cuatro indicadores o carencias que sufren las personas. De las variables que se consideran para el IPM, solo se considera el acceso a internet, la disponibilidad de agua potable y la tenencia de servicio sanitario, juntas con una ponderación de 15% dentro del índice.

Con respecto a los hogares que registran simultáneamente los tres tipos de pobreza, de acuerdo a los datos de la tabla 12, se pueden identificar un total de 28.517 hogares, correspondiente a 95.022 personas, donde el mayor porcentaje se ubica en la zona rural y en específico, las regiones Central, Huetar Atlántica y Huetar Norte, las cuales en forma conjunta representan 73,4% de los hogares con estas características.

Para terminar y tomando en cuenta los resultados hasta ahora obtenidos, se pueden resaltar algunas propiedades que cumple el IPM como indicador de medida de la pobreza. Como bien señalan Gradín y del Río (2001), dado el gran número de índices de pobreza existente en la literatura teórica son necesarios criterios que permitan priorizar su utilización en un estudio empírico concreto.

a) Exhaustividad: el IPM aprovecha al máximo y de forma útil la información suministrada por las variables seleccionadas.

b) Enfoque: el valor del IPM de un hogar es independiente de los niveles de dotación de bienes y servicios de los demás hogares.

c) Unicidad: para cada hogar considerado en la muestra se garantiza un IPM.

d) Simetría: los hogares con una idéntica dotación de bienes y servicios reciben un mismo valor IPM.

Tabla 12

*Cantidad y porcentaje de hogares pobres por zona y región. Pobreza IPM + pobreza por insuficiencia de ingresos + pobreza multidimensional*

Zona y región	Cantidad	Porcentaje
Costa Rica	28.517	100
Zona		
Urbano	12.440	43,6
Rural	16.077	56,4
Región		
Central	8.147	28,6
Chorotega	2.717	9,5
Pacífico Central	2.230	7,8
Brunca	2.634	9,2
Huetar Atlántica	6.053	21,2
Huetar Norte	6.736	23,6

Fuente. Porras Solís, 2018

e) Progresividad: los valores del IPM reflejan una tendencia creciente conforme la dotación de bienes y servicios de los hogares aumenta.

f) Descomponibilidad aditiva: el IPM de una muestra es equivalente a la suma ponderada de los índices de cada uno de los  $k$  grupos de que está compuesta.

g) Continuidad: si se produce un cambio muy pequeño en la dotación de bienes y servicios de los hogares, entonces se produce una variación también pequeña en el correspondiente valor del IPM.

h) Invariancia ante cambios de escala: el indicador IPM es invariante a un cambio de origen y/o de escala en las unidades en que venga expresados.

i) Discriminación: los valores que toma el IPM para el conjunto de hogares de la muestra manifiestan un recorrido o intervalo de variación suficientemente amplio, que permiten diferenciar la situación de deterioro o bienestar de los hogares.

j) Transitividad: si  $a, b$  y  $c$  son tres situaciones distintas de pobreza para tres hogares, y si  $IPM_a, IPM_b$  y  $IPM_c$  son los valores del indicador correspondiente, se verifica que  $IPM_a > IPM_b$  y  $IPM_b > IPM_c$  implica que  $IPM_a > IPM_c$ .

## Conclusiones

En definitiva, la pobreza es un fenómeno socioeconómico complejo, con múltiples causas y consecuencias, del cual no se posee un único indicador que dé cuenta de sus diversas dimensiones. En el caso de Costa Rica, durante muchos años las mediciones de pobreza se han concentrado en captar una de estas dimensiones a través de los ingresos monetarios disponibles y el gasto en consumo. Por tanto, el análisis de los elementos de equipamiento de los hogares resulta de gran interés, ya que permite abordar el problema de la medición de la pobreza desde

otra perspectiva, mediante un indicador no monetario.

Las diversas técnicas estadísticas multivariantes utilizadas permitió cumplir con el principal objetivo del trabajo: elaborar un índice para la medición de la pobreza a partir de los elementos de equipamiento de los hogares. En primer lugar, con la medida de distancia DPM se logra reducir la dimensionalidad del conjunto de las variables seleccionadas, obteniéndose una ordenación inicial de los hogares, donde aquellos con valores bajos carecen de prácticamente todo el equipamiento, mientras que en el extremo, aparecen los hogares con mayor equipamiento.

En función de este resultado, se aplicó un análisis discriminante, obteniéndose una función canónica discriminante, cuyos puntajes son utilizados como un índice para medir y ordenar, la situación de la pobreza de los hogares.

Mediante el análisis de clúster se consiguió establecer un grupo de bienes y servicios de primera necesidad o básicos en los hogares (dotación mínima), a saber: suministro de agua potable, tenencia de servicio sanitario, tenencia de baño, suministro de electricidad, refrigeradora, televisor y teléfono celular. Esta dotación mínima permitió obtener un valor crítico para establecer una frontera de distinción entre hogares considerados como pobres y no pobres. De esta manera, los hogares con un IPM menor a dicho valor, se clasificaron como hogares pobres y los que registraban un valor más alto, fueron clasificados como hogares no pobres. Dicha línea de pobreza permitió determinar que 6,23 % de los hogares se encuentran en condición de pobreza.

La aplicación empírica del índice utilizando datos de la ENAHO 2016 permitió obtener importantes conclusiones acerca de la incidencia de la pobreza a nivel territorial. Al respecto, la pobreza afecta en mayor medida a los hogares



ubicados en la zona rural y en particular a los situados en las afueras y alejadas del valle central del país. Precisamente, las regiones Huetar Atlántica y Huetar Norte, son las que presentan un mayor porcentaje de hogares en condiciones de pobreza. Es importante señalar que el IPM también puede ser utilizado para estudiar su correlación (o cruce) con otro tipo de variables diferentes a las territoriales que incluye la ENAHO 2016, tales como las sociodemográficas, económicas, empleo, etc.

Desde la perspectiva de los ingresos, se determinó que un poco más del 70% de estos hogares se ubican en los dos quintiles más bajos de ingresos per cápita del país. De la misma forma se determinó que un 43,2% de los hogares pobres se encuentran en condición de pobreza por insuficiencia de ingresos, mientras que, en términos de pobreza multidimensional, significa 53,7%. En cuanto a los hogares que tienen simultáneamente los tres tipos de pobreza, estos se ubican en mayor porcentaje en la zona y en las regiones fuera del valle central, principalmente, las regiones Huetar Atlántica y Huetar Norte.

Para concluir, el indicador de pobreza IPM reúne una serie de propiedades, que, en efecto, permiten abordar el problema de la medición de la pobreza a partir de los elementos de equipamiento de los hogares. Algunas de las propiedades como la descomponibilidad aditiva y la discriminación, resultan ser fundamentales para el análisis de la incidencia territorial del fenómeno.

## Referencias

- Céspedes, V. y Jiménez, R. (2006). *Pobreza en Costa Rica*. III Jornada anual de la Academia de Centroamérica. San José, Costa Rica. Recuperado de <http://enlaceacademico.ucr.ac.cr/sites/default/files/publicaciones/Librojornada3.pdf>
- Díaz, M., González, A, Henao, A. (2013). *Introducción al análisis estadístico multivariado aplicado*. Editorial Universidad del Norte. Barranquilla, Colombia.
- Fournier, M. y Garner L. (1990). Fórmula abreviada para la transformación de puntajes de variables a una misma escala. *Revista Latinoamericana de Psicología* 22(3), 421-426. Recuperado de <http://www.redalyc.org/pdf/805/80522305.pdf>
- García, G., Sánchez, P. y Vaamonde A. (2007). *A pobreza en Galicia. Medición a través do equipamiento dos fogares*. Sotelo Blanco Edicions. Santiago de Compostela, España.
- Gradin, C. y del Río, C. (2001). *Desigualdad, Pobreza y Polarización en la Distribución de la Renta en Galicia*. Instituto de Estudios Económicos de Galicia-Pedro Barrié de la Maza. A Coruna, España. Recuperado de [http://pmayobre.webs.uvigo.es/06/arch/profesorado/olga\\_sanchez/desigualdad.pdf](http://pmayobre.webs.uvigo.es/06/arch/profesorado/olga_sanchez/desigualdad.pdf)
- Hernández, O. (2013). *Temas de análisis estadístico multivariante*. Segunda edición. Editorial UCR. San José, Costa Rica.
- INE (2007). *La pobreza y su medición. Presentación de diversos métodos de obtención de medidas de pobreza*. Madrid, España. Recuperado de <http://www.ine.es/daco/daco42/sociales/pobreza.pdf>
- INEC (2010). *Actualización metodológica para la medición del empleo y la pobreza*. Recuperado de [http://www.inec.go.cr/sites/default/files/documentos/pobreza\\_y\\_presupuesto\\_de\\_hogares/pobreza/metodologias/documentos\\_metodologicos/mepobrezaena-ho2010-02.pdf](http://www.inec.go.cr/sites/default/files/documentos/pobreza_y_presupuesto_de_hogares/pobreza/metodologias/documentos_metodologicos/mepobrezaena-ho2010-02.pdf)
- INEC (2015). *Índice de pobreza multidimensional: metodología*. Recuperado de [http://www.inec.go.cr/sites/default/files/documentos/pobreza\\_y\\_presupuesto\\_de\\_hogares/pobreza/metodologias/mepobrezaena-ho2015-01.pdf](http://www.inec.go.cr/sites/default/files/documentos/pobreza_y_presupuesto_de_hogares/pobreza/metodologias/mepobrezaena-ho2015-01.pdf)

- INEC (2016). *Encuesta Nacional de Hogares (ENAHO) 2016. Resultados generales*. San José, Costa Rica. Recuperado de [http://www.inec.go.cr/sites/default/files/documentos-biblioteca-virtual/reenaho2016\\_0.pdf](http://www.inec.go.cr/sites/default/files/documentos-biblioteca-virtual/reenaho2016_0.pdf)
- Larraín, F. (2008). Cuatro millones de pobres en Chile: Actualizando la Línea de Pobreza. *Estudios Públicos* (109), 101-148.
- Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M. y Hornik, K. (2013). *cluster: Cluster Analysis Basics and Extensions*. R package version 1.14.4.
- Pena, J. (1977). *Problemas de medición del bienestar y conceptos afines, (una aplicación al caso español)*. INE. Madrid, España.
- R Core Team (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Townsend, M. (1993). La conceptualización de la pobreza. *Revista Comercio Exterior* 53(6), 445-452.
- Venables, W. y Ripley, B. (2002). *Modern Applied Statistics with S*. (Fourth Edition). Springer, New York, USA.
- Zarzosa, P. (2009). Estimación de la pobreza en las comunidades autónomas españolas, mediante la distancia  $DP_2$  de Pena. *Estudios de Economía Aplicada* (27), 397- 416.

Recibido: 28 de febrero de 2018

Aceptado: 14 de abril de 2018