

---

# Indentificación de una tipología de pobreza multidimensional a través del enfoque de cluster probabilístico

---

Laura Nalbarte

Silvia Altmark

Fernando Massa

**Serie Documentos de Trabajo**

Nº 2/22

1er semestre, 2022

ISSN: 1688-6453

**Universidad de la República.**

**Facultad de Ciencias Económicas y de Administración,**

**Instituto de Estadística (IESTA)**

**Montevideo, Uruguay.**



Esta obra está bajo una Licencia Creative Commons Atribución - NoComercial - CompartirIgual 4.0 Internacional.

Forma de citación sugerida para este documento:

Laura Nalbarte, Silvia Altmark, Fernando Massa (2022). *Indentificación de una tipología de pobreza multidimensional a través del enfoque de cluster probabilístico* (Serie Documentos de Trabajo; Nº2/22). Montevideo: Universidad de la República. Facultad de Ciencias Económicas y de Administración, Instituto de Estadística.  
<https://www.colibri.udelar.edu.uy/jspui/handle/20.500.12008/10518>

# Identificación de tipología de pobreza multidimensional a través del enfoque de cluster probabilístico

Laura Nalbarte <sup>1</sup> Silvia Altmark <sup>2</sup> Fernando Massa <sup>3</sup>;

*Instituto de Estadística*

*Facultad de Ciencias Económicas y de Administración*

*Universidad de la República*<sup>4</sup>

## Resumen

La medición y caracterización de la pobreza constituye un tema de interés social que estadísticamente puede ser abordado desde distintas perspectivas. Mediante la utilización con técnicas de Análisis Factorial y/o Análisis de Cluster se definieron particiones representativas de hogares con diferente grado de vulnerabilidad.

El objetivo de la investigación es la medición multidimensional de la pobreza y la construcción de tipologías de hogares (pobres, vulnerables y no pobres) a partir de datos relevados por la Encuesta Continua de Hogares (ECH) del Instituto Nacional de Estadística de Uruguay (INE), entre los años 2008 y 2012. En este documento se presentan los resultados correspondientes a un proceso de clustering, basado en una mezcla probabilística de distribuciones Bernoulli. El mismo asume la presencia de una variable categórica no observable que dictamina la pertenencia de los individuos a los grupos. Los parámetros se estiman por medio del algoritmo EM y los individuos se clasifican utilizando probabilidades “a posteriori”. Las variables consideradas en el análisis refieren, tanto a características de la vivienda (materiales de techos, paredes y pisos), como a características del hogar y de sus integrantes. Respecto al hogar, se tiene en cuenta: la composición del mismo, el clima educativo, la cantidad de niños menores de 6 años, la tenencia y acceso a tecnologías de la información y la comunicación (TICS), así como la tenencia de bienes de confort. En lo que refiere a las personas, se considera la situación laboral del jefe, el ingreso y su etnia. Los resultados indican la existencia de 3 ó 4 grupos que determinan un gradiente de vulnerabilidad, desde los hogares más críticos (pobres-más vulnerables), hasta los que se encuentran en mejores condiciones (menos vulnerables).

---

<sup>1</sup> *email:* [laura.nalbarte@fcea.edu.uy](mailto:laura.nalbarte@fcea.edu.uy)

<sup>2</sup> *email:* [salt@iesta.edu.uy](mailto:salt@iesta.edu.uy), ORCID: <https://orcid.org/0000-0002-3291-7820>

<sup>3</sup> *email:* [fernando.massa@fcea.edu.uy](mailto:fernando.massa@fcea.edu.uy), ORCID: <https://orcid.org/0000-0002-2922-4097/>

<sup>4</sup> Este trabajo forma parte de la producción póstuma de la Prof. Laura Nalbarte

**Palabras clave:** Pobreza multidimensional, Cluster probabilístico.

**CÓDIGOS JEL:** C10, C18, C38, C83.

**CLASIFICACIÓN MATEMÁTICA POR TEMAS MSC2020:** 62P25, 62H25, 62H30, 62H86 .

### **ABSTRACT**

The measurement and characterization of poverty is a topic of social interest that is statistically it can be approached from different perspectives. By using using Factorial Analysis and/or Cluster Analysis techniques, partitions were defined trough households with different degrees of vulnerability. The objective of the research is the multidimensional measurement of poverty and construction tion of typologies of households (poor, vulnerable and non-poor) from data surveyed by the Encuesta Continua de Hogares -ECH (Continuous Household Survey) of the Instituto Nacional de Estadística de Uruguay INE (National Institute of Statistics of Uruguay). between 2008 and 2012. This document presents the results corresponding to a clustering process, based on a probabilistic mixture of distributions Bernoulli butions. It assumes the presence of an unobservable categorical variable that dictates the membership of individuals to groups. The parameters are estimated by means of the EM algorithm and the individuals are classified using a posteriori probabilities. The variables considered in the analysis refer both to characteristics of the dwelling (ceiling, wall and floor materials), as well as characteristics of the home and its great. Regarding the home, the following are taken into account: its composition, the educational climate, the number of children under 6 years of age, possession and access to information technologies and communication (TICS), as well as the possession of comfort goods. In regards to people, the employment status of the boss, income and ethnicity are considered. The results indicate the existence of 3 or 4 groups that determine a gradient of vulnerability, from the most critical households (poor-most vulnerable), to those are in better conditions (less vulnerable).

**Key words:** Multidimensional poverty, Probabilistic cluster..

**JEL CODES:** C10, C18, C38, C8 .

**MATHEMATICS SUBJECT CLASSIFICATION MSC2020:**62P25, 62H25, 62H30, 62H86.

## 1. Introducción

El tema de la pobreza en Uruguay, como en toda América Latina, es relevante desde el punto de vista social. La medición y caracterización de la pobreza constituye un tema de interés social, que puede ser abordado estadísticamente desde distintas perspectivas: elaboración de índices, análisis factorial o construcción de grupos, entre otros.

En Uruguay el cálculo oficial de la pobreza lo realiza el Instituto Nacional de Estadística (INE) mediante el método del ingreso. Las fuentes de información utilizadas son la Encuesta Nacional de Gastos e Ingresos 2005-2006 (ENGIH) y las Encuestas Continuas de Hogares (ECH) realizadas anualmente. En esta línea, determina si un hogar es pobre comparando su ingreso con la línea de pobreza (LP), según la metodología indicada en el documento “Estimación de la pobreza por el método del ingreso 2013” (INE, 2013).

Este trabajo tiene por objetivos la medición de la pobreza multidimensional (PM) en Uruguay entre los años 2008 y 2012 a través de la construcción de tipologías de hogares (pobres, vulnerables y no pobres), utilizando datos provenientes de las ECH del INE.

Con respecto al ciclo de vida de este trabajo se deja constancia que un primer avance de esta investigación fue presentado como preprint en el portal <https://doi.org/10.31235/osf.io/nv962>. Este trabajo es parte de la obra póstuma de la Prof. Laura Nalbarte.

## 2. Antecedentes

Algunos de los antecedentes más recientes en materia de la cuantificación PM son los que se enumeran a continuación.

En las “Notas Técnicas del Informe sobre desarrollo humano 2010” de Naciones Unidas (Klugman, 2010), se plantea un índice de Pobreza Multidimensional (IPM) para Colombia que identifica privaciones individuales en materia de salud, educación y nivel de vida. Usando micro datos de encuestas de hogares, clasifica a cada persona de un hogar como pobre o no, dependiendo de la cantidad de privaciones a que está sometida su familia. Finalmente, los datos se agregan en la medición nacional de pobreza. El IPM se basa en tres componentes: Salud, Educación y Nivel de vida, cada uno de los cuales se mide a través de al menos 2 indicadores, con los cuales se determinan dos mediciones de pobreza: la intensidad y la tasa de incidencia. La intensidad de la pobreza refleja la proporción de los

indicadores de los componentes ponderados donde, en promedio, los pobres están sujetos a privaciones. La tasa de incidencia es la proporción de la población multidimensionalmente pobre. El IPM evalúa cinco dimensiones donde los hogares colombianos podrían estar en privación. Ambas medidas son complementarias y no excluyentes, ya que son mediciones desde ópticas diferentes del mismo fenómeno.

Para el cálculo de la pobreza monetaria se definió una nueva metodología. El IPM (Alkire y Foster, 2007) es un índice desarrollado por la Universidad de Oxford que fue adaptado para Colombia, incluyendo cinco dimensiones: condiciones educativas del hogar; condiciones de la niñez y la juventud; salud; trabajo y servicios públicos y condiciones de la vivienda. Estas cinco dimensiones son evaluadas a través de 15 indicadores obtenidos de la Encuesta Nacional de Calidad de Vida; aquellos hogares que tengan “privación” en por lo menos el 33 % son considerados pobres.

La “Medición ampliada del bienestar” (Wolff *et al.*, 2012), desarrollada por el Levy Economics Institute, parte de la idea de que los ingresos (laborales o no) son la medida estándar del bienestar, ya que indican la capacidad de compra de los hogares, es decir, son una buena aproximación al consumo de los hogares. Sin embargo su índice incluye información de transferencias gubernamentales, consumo público e ingreso proveniente de la riqueza e impuestos entre otros. En una aplicación se compara la mediana de este índice en Canadá con respecto a los Estados Unidos llegando a la conclusión de que la brecha entre ambos países en 2007 fue de 7 % en 2005, mientras que considerando el ingreso bruto, la diferencia entre las medianas es de 3 %.

En el caso de Uruguay, el antecedente más cercano en el tiempo es el referente a la aplicación de la metodología del Consejo Nacional de Evaluación de la Política de Desarrollo Social de México-CONEVAL (CONEVAL, 2010) realizada por el MIDES en el 2012. En dicho documento se calculan Indicadores asociados a las dimensiones de Bienestar y Derechos Sociales. En el caso de la medición del Bienestar, se consideran las Líneas de Pobreza e Indigencia, mientras que, para los indicadores asociados a los Derechos sociales se utilizan cifras vinculados a carencias en educación, al acceso a servicios de salud, el acceso a seguridad social, entre otras.

Otro antecedente vinculado a la cuantificación de la PM en el Uruguay, es el “Atlas Sociodemográfico y de la desigualdad en el Uruguay” (Calvo *et al.*, 2011) que presenta resultados referidos a las Necesidades Básicas Insatisfechas (NBI) calculadas a partir de los

datos del censo de año 2011. Dicho trabajo propone una revisión de las seis dimensiones que caracterizan las NBI, los indicadores que las componen y los umbrales críticos que determinan la ausencia/presencia de cada NBI. En este trabajo se plantean las diferencias que existen a nivel geográfico, etario y según ascendencia concluyendo que “... *La situación de la población respecto al nivel de carencias críticas observadas varía significativamente al considerar la edad, lugar de residencia, comportamiento reproductivo, condición migratoria, ascendencia étnico-racial y limitaciones de las personas, encontrándose en muchos casos distancias importantes en los valores de estos grupos con respecto a los promedios, lo cual evidencia situaciones de desigualdad...*”.

Más atrás en el tiempo se puede encontrar el enfoque de Katzman (Katzman, 1989). En el mismo se combinan las ideas de pobreza según el método del ingreso y según las NBI. En dicho trabajo, Katzman plantea las cuatro situaciones (véase Figura 1) que surgen de cruzar ambos criterios.

De esta manera, diferencia la situación (cuantificada a través de las NBI) de los pobres

	NBS	NBI
Hogares por encima de la LP	Hogares en condiciones de integración social	Pobreza inercial
Hogares por debajo de la LP	Pobreza reciente	Pobreza crónica

Figura 1: Enfoque de Katzman

en tres subgrupos: pobres recientes, inerciales y crónicos.

### 3. Metodología

#### 3.1. Datos

Las bases de datos utilizadas en este trabajo son las correspondientes a las ECH llevadas a cabo por el INE entre los años 2008 y 2012. A partir de 2006 las encuestas tienen cobertura nacional. Las bases de datos son el resultado de una muestra de diseño complejo llevada a cabo por el INE en cada año. Las variables tenidas en cuenta para este trabajo

describen características tanto del hogar (composición, clima educativo, acceso a tecnologías de la información (TICS)), la vivienda (materiales de techo, paredes y piso) y sus integrantes (bienes de confort, situación laboral del jefe). La información es analizada a nivel de hogares, trabajando con los datos expandidos.

Tanto el pre-procesamiento como el análisis de los datos fue llevado a cabo en el software de libre distribución R (R Core Team, 2013).

### 3.2. Métodos Aplicados

La estrategia de análisis llevada a cabo en este trabajo fue la de crear grupos mediante un modelo probabilístico, de manera que el insumo final no fuese la pertenencia de cada hogar a un grupo específico sino el grado de “afinidad” con los elementos de cada uno de los grupos.

El análisis de grupos (o clusters) es una técnica estadística multivariada que consiste en construir una partición del espacio en el que se encuentra un conjunto de individuos, de forma tal que los grupos constituídos por estos individuos sean lo más homogéneos “dentro” de sí y lo más heterogéneos posible “entre” ellos. Cabe agregar que el análisis de grupos no constituye un algoritmo de por sí sino que puede ser resuelto por diversos algoritmos.

En cuanto al análisis de cluster probabilístico, este documento sigue la metodología propuesta en (Moustaki y Papageorgiou, 2005). Se utiliza un modelo de clase latente donde se asume la existencia de una variable aleatoria  $Z$  (no observable) que consta de  $G$  clases. En esta etapa, el objetivo es realizar predicciones sobre el valor de  $Z$  utilizando un conjunto de variables explicativas.

Sea  $(x_1, x_2, \dots, x_p)$  un vector de  $p$  variables con distribución Bernoulli con parámetro  $\phi_{jk}$  donde el subíndice  $k$  hace referencia a cada una de las variables y  $j$  a cada uno de los grupos. Para completar la especificación del modelo es necesario definir las probabilidades de mezcla  $\tau_j$ , las mismas suelen llamarse probabilidades a priori. La distribución conjunta de las variables observadas está dada por la siguiente mezcla finita probabilística:



$$f(X, Z) = \prod_{j=1}^G \prod_{k=1}^p [\tau_j \phi_{jk}^{x_{ik}} (1 - \phi_{jk})^{1-x_{ik}}]^{1(Z_i=j)}$$

Es necesario incluir la variable aleatoria  $I(Z_i = j)$  debido a que el problema que intenta solucionar el análisis de grupos es que no se sabe a qué grupo pertenece cada observación. De esta manera, es posible plantear el problema de la asignación de individuos a grupos mediante la distribución de la variable  $Z$ . Los parámetros del modelo se estiman iterativamente a través del algoritmo *EM* (Dempster *et al.*, 1977). Para este caso, cada iteración de dicho algoritmo, consta de las siguientes etapas:

- paso E)  $Q(\theta|\theta^t) = \mathbb{E}_{Z|X, \theta^t} \mathcal{L}(\theta|X, Z)$
- paso M)  $\theta^{t+1} = \max_{\theta} Q(\theta|\theta^t)$

El valor de la log-verosimilitud de una muestra de  $n$  individuos está dada por la siguiente expresión:

$$\mathcal{L}(\theta|X, Z) = \sum_{i=1}^n \sum_{j=1}^G 1(Z_i = j) \sum_{k=1}^p [\log \tau_j + x_{ik} \log \phi_{jk} + (1 - x_{ik}) \log (1 - \phi_{jk})]$$

Este proceso iterativo se repite hasta que el cambio en  $Q(\theta|\theta_t)$  sea menor que cierta tolerancia especificada previamente.

Uno de los principales inconvenientes de este algoritmo es su dependencia del valor inicial del vector de parámetros. Para hacer frente a este problema, en este trabajo se optó por inicializar el algoritmo con distintos valores iniciales y seleccionar aquel modelo cuyo valor de la  $Q(\theta|\theta_t)$  sea mayor. Algunas de las ventajas de este enfoque son las siguientes:

- Dada la naturaleza paramétrica del modelo, se puede seleccionar el número de grupos utilizando criterios de información como el Bayesian Information Criterion (BIC) (Schwarz, 1978) o el Akaike Information Criterion (AIC) (Akaike, 1974).
- A través de la distribución a posteriori de la variable aleatoria  $Z$  (condicional a los parámetros y a las variables explicativas), se puede definir el grado de pertenencia de cada individuo a cada grupo.

$$P(Z_i = j) = \frac{\tau_j \prod_{k=1}^G \phi_{jk}^{x_{ik}} (1 - \phi_{jk})^{1-x_{ik}}}{\sum_{j=1}^G \tau_j \prod_{k=1}^G \phi_{jk}^{x_{ik}} (1 - \phi_{jk})^{1-x_{ik}}}$$

- Mediante los parámetros del modelo se pueden calcular frecuencias esperadas que, comparadas con las frecuencias observadas, indiquen el grado de ajuste del modelo.

## 4. Resultados

### 4.1. Descripción preliminar

Como punto de partida, se realiza una descripción de los hogares pobres en Uruguay, según las siguientes variables: sexo del jefe del hogar, cantidad de niños menores de 6 años, tipo de hogar (compuesto, extendido, nuclear monoparental, nuclear completo, unipersonal) y ascendencia del jefe del hogar (blanco, afrodescendiente, otra). La tabla 1 presenta distribución de los hogares según estas variables.

Puede observarse cómo todas las medidas se mantienen relativamente estables dentro del período de estudio. Estas medidas son utilizadas como medición de base para comparar los resultados obtenidos luego de la generación de los grupos.

Adicionalmente, se calcularon las NBI, según la metodología del año 2011. Los resultados se exponen en la tabla 2.

En este caso se puede observar como la proporción de hogares libres de NBI aumenta a lo largo del período. Pese a que toda la proporción de hogares carentes de cada necesidad presenta una tendencia a la baja, el aumento en el número de hogares libres de NBI parece deberse (mayoritariamente) al descenso sostenido de la NBI “artículos de confort”.

### 4.2. Cluster Probabilístico

Dada la naturaleza binaria de las variables, se utilizó el método de cluster probabilístico para la construcción de grupos. Éste tiene la ventaja sobre los métodos convencionales de no depender de la elección de una distancia. Esto debe ser destacado debido a que,

Tabla 1: Descripción inicial de los hogares

	2008	2009	2010	2011	2012
Sexo del jefe del hogar					
Femenino	35,8 %	36,4 %	37,2 %	35,6 %	40,5 %
Masculino	64,2 %	63,6 %	62,8 %	64,4 %	59,5 %
Número de niños					
0	79,8 %	79,5 %	79,8 %	78,9 %	79,3 %
1	14,8 %	15,2 %	15,1 %	16,0 %	15,4 %
2	4,4 %	4,4 %	4,1 %	4,3 %	4,5 %
3 ó más	1,0 %	0,9 %	1,0 %	0,9 %	0,8 %
tipo de hogar					
Unipersonal	21,3 %	21,8 %	21,9 %	20,9 %	1,1 %
Nuclear completo	50,5 %	50,6 %	50,8 %	50,7 %	20,0 %
Nuclear monoparental	11,7 %	11,7 %	11,8 %	12,1 %	51,0 %
Extendido	14,6 %	13,9 %	13,7 %	14,5 %	12,4 %
Compuesto	1,9 %	2,0 %	1,9 %	1,9 %	15,5 %
Ascendencia de jefe del hogar					
Otra	0,2 %	0,3 %	0,2 %	0,2 %	0,3 %
Afrodescendiente	8,2 %	7,7 %	7,9 %	7,6 %	8,8 %
Blanca	91,6 %	92,1 %	91,9 %	92,1 %	90,9 %

Fuente: Elaboración propia en base a datos del INE

a diferencia de los datos cuantitativos con los que comúnmente se opera en este tipo de análisis, las distancias basadas en variables binarias presentan características no deseables, como por ejemplo, que distintas parejas de configuraciones (que indican situaciones sumamente diferentes) presentan distancias iguales. Otra característica positiva de este método es que no tiene “pérdida de información”. Se menciona esto debido a que una alternativa para trabajar con distancias definidas sobre variables cuantitativas es proyectar los datos utilizando técnicas de análisis factorial, método que, como es bien sabido, trata de construir variables que eliminan información “redundante” en el conjunto de datos. Sin embargo, el criterio de redundante es subjetivo y queda a criterio del usuario.

En primer lugar se puso en marcha el algoritmo para decidir el número de grupos a considerar para los análisis posteriores. De esta manera, en cada año se estimaron los parámetros de la mezcla probabilística teniendo en cuenta situaciones de 2, 3, 4 y 5 grupos. En la tabla 3 se presentan los resultados de esta etapa, los cuales sugieren que trabajar con tres grupos parece adecuado.

Tabla 2: Proporción de hogares según NBI

	2008	2009	2010	2011	2012
NBI - Indicadores					
Materialidad	0,8 %	0,7 %	0,7 %	0,6 %	0,3 %
Espacio habitable	4,0 %	3,5 %	3,5 %	3,0 %	3,1 %
Espacio cocinar	9,1 %	8,5 %	8,3 %	7,3 %	6,0 %
Agua potable	9,9 %	5,7 %	5,0 %	4,6 %	3,4 %
Servicio higiénico	4,8 %	4,3 %	4,0 %	4,3 %	3,6 %
Energía eléctrica	1,2 %	1,1 %	0,9 %	0,9 %	0,5 %
Artículos confort	16,1 %	15,7 %	14,3 %	12,4 %	10,4 %
Educación	5,0 %	4,4 %	4,6 %	5,0 %	4,5 %
N° de NBI					
Ninguna	76,6 %	76,6 %	78,0 %	79,3 %	81,7 %
Al menos una	23,4 %	23,4 %	22,0 %	20,7 %	18,3 %

Fuente: Elaboración propia en base a datos del INE

Tabla 3: BIC según N° de grupos por año

N° grupos	2008	2009	2010	2011	2012
5	971030	830582	869146	775554	740637
4	1004964	846863	839685	757144	755885
3	908782	816973	800011	806291	705215
2	912543	834590	1000148	859540	831528

Fuente: Elaboración propia

Se conformaron tres grupos siguiendo la metodología de análisis de grupos basado en un modelo probabilístico. La Figura 2 presenta el tamaño de cada grupo dentro de la mezcla probabilística para cada uno de los años.

Se observa una leve tendencia a la disminución del grupo catalogado como “bajo” en favor del aumento del grupo “medio-alto”.

El siguiente paso fue la caracterización de los grupos en función de las proporciones estimadas para cada variable en la mezcla probabilística. De esta manera se concluyó que uno de los grupos se asoció a hogares más carenciados, otro se conformó a partir de hogares en

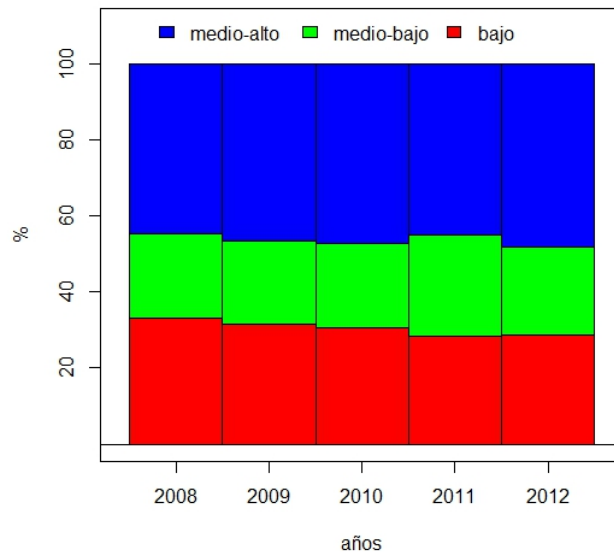
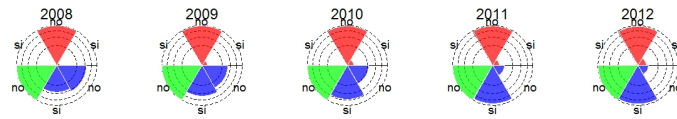
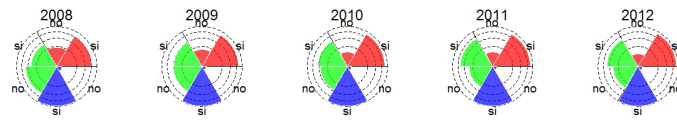


Figura 2: Tamaño relativo de los grupos a través de los años

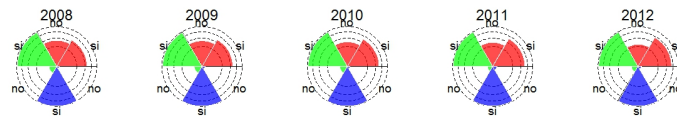
mejores condiciones y el restante se formó con hogares en una situación intermedia. En las Figuras 3a-3e se muestra la evolución de algunas de las variables utilizadas para construir los grupos: tenencia de Internet, de celular, disponibilidad de agua caliente, tenencia de lavarropa y número de empleos del jefe del hogar.



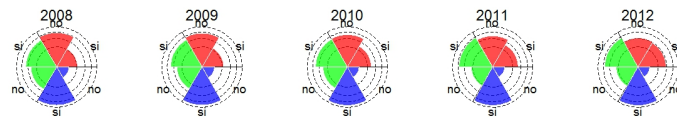
(a) Tenencia de internet según grupo para cada año



(b) Tenencia de celular según grupo para cada año



(c) Disponibilidad de agua caliente según grupo para cada año



(d) Tenencia de lavarropa según grupo para cada año



(e) N° de empleos del jefe del hogar según grupo para cada año

Figura 3: Distribución de variables de confort según grupo y año

En el caso de la tenencia de Internet, se observa un aumento sostenido en el grupo catalogado como “medio-alto” (azul) y un leve incremento en el grupo “bajo” (rojo). En cuanto a la tenencia de celular, se observa que para el 2012 casi la totalidad del grupo “bajo” está conformado por hogares que tienen celular. El grupo “medio-bajo” (verde) también muestra un ascenso marcado mientras que los hogares del grupo “alto” ya contaban con celulares desde el 2008 o antes.

La variable agua caliente se construyó indicando que el hogar cuenta con calefón o calentador. De la misma manera que en la variable anterior, los hogares del grupo “alto” ya disponen de este bien desde el 2008 o antes. En el caso del grupo “medio” se ve que casi todos los hogares cuentan con agua caliente, en tanto que los del grupo “bajo” evidencian una leve evolución temporal hacia una condición más favorable.

La tenencia de lavarropas en el grupo “alto” se da en casi todos los casos. En el grupo “medio” se nota un aumento sostenido a lo largo del periodo de los hogares que cuentan con este bien. En el grupo “bajo” también se ve un aumento pese a que comienza en una situación más desfavorable que la del grupo “medio”. Por último, se muestra la situación del número de empleos de jefe del hogar. Pese a que no hay una gran variación temporal, sí se puede ver que el multiempleo se da en los grupos “alto” y “bajo”.

La siguiente instancia fue la de validar los grupos mediante variables no utilizadas en su construcción. En esta etapa se tuvieron en cuenta el ingreso per cápita, los quintiles de ingreso, el clima educativo, el tipo de hogar, el número de niños menores a 6 años, el sexo y la ascendencia del jefe del hogar. Considerando en cada hogar las personas de 18 años o más, se definió la variable “años de educación promedio del hogar”. En la Figura 4 se puede apreciar cómo el grupo “medio-alto” presenta situaciones más favorables a través de los años, en tanto perciben ingresos mayores y presentan situaciones educativas más propicias.

En una siguiente etapa, a modo de profundizar en las diferencias en cuanto el ingreso entre grupos, se investigó la composición de los mismos según los quintiles de ingreso. La Figura 5 presenta la situación al comienzo y al final del período.

Se puede ver que la situación no presenta variaciones claras a lo largo del período; sin embargo, pueden observarse distribuciones claramente distintas entre grupos. Casi la mitad del grupo “bajo” está compuesto por hogares en el primer quintil de ingresos, mientras que en los otros dos grupos el primer quintil representa menos del 10%.

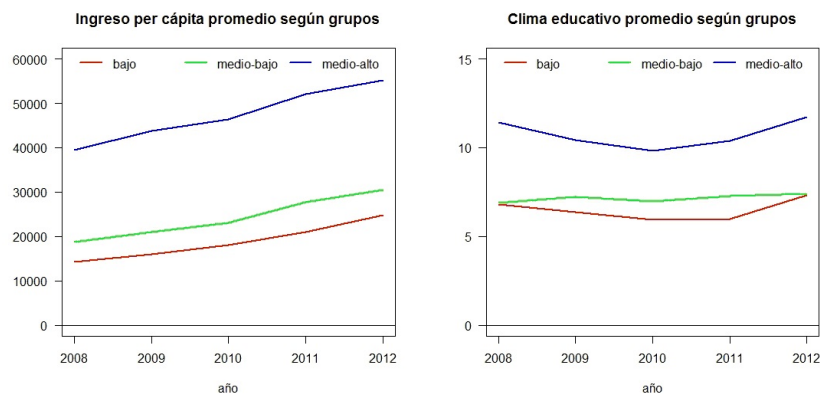


Figura 4: Distribución del clima educativo e ingreso per cápita según grupo par cada año

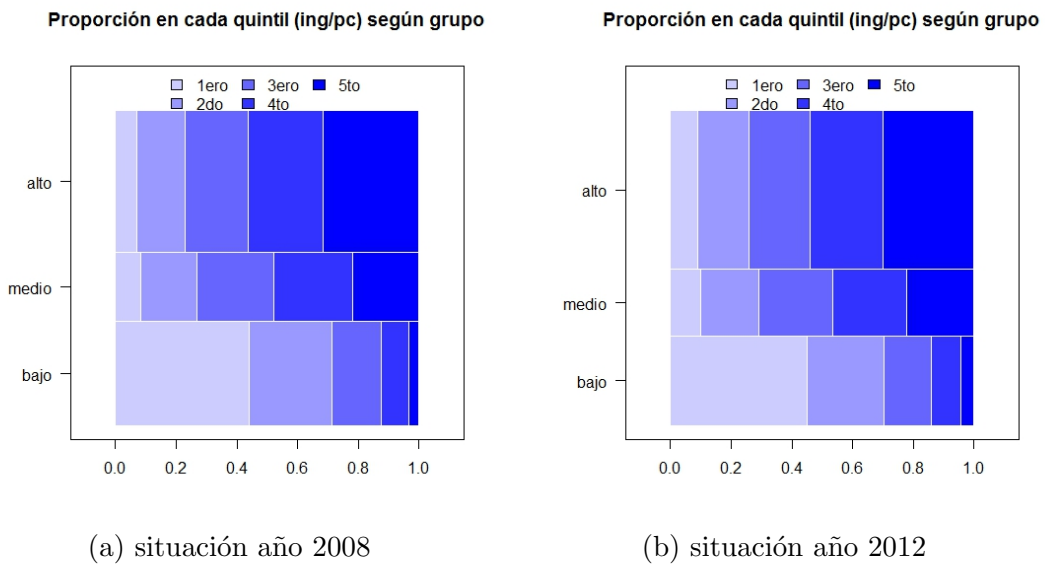
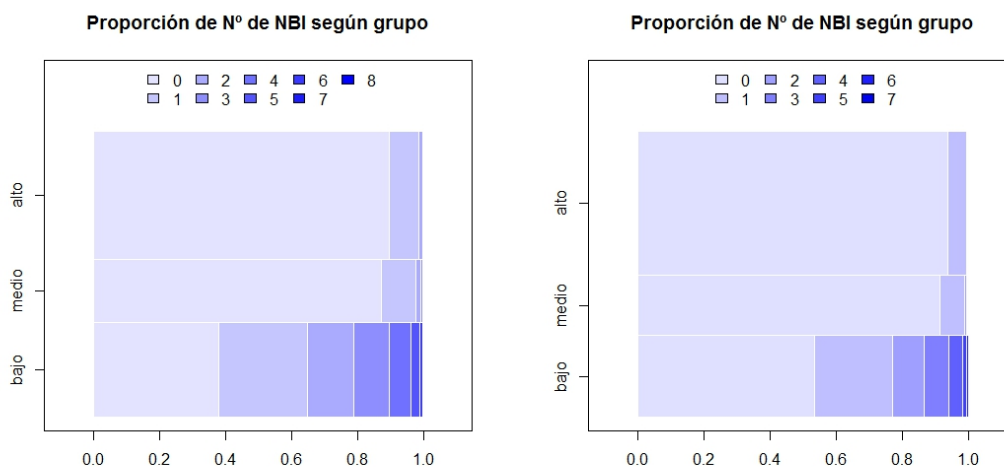


Figura 5: Quintiles de ingreso según grupos



Un análisis similar se llevó a cabo con el número de NBI por grupo (véase Figura ??). La situación es similar a la analizada previamente: mientras que los grupos menos vulnerables sólo presentan una, o a lo sumo dos NBI, el grupo catalogado como “bajo” presenta hogares en todas las categorías. No obstante, puede observarse una evolución dentro de este grupo, ya que en 2008, poco menos del 40 % de los hogares que lo componen estaba libre de NBI, mientras que en 2012, esta cifra asciende hasta casi el 60 %.



(a) Número de NBI por grupo en 2008      (b) Número de NBI por grupo en 2012

Figura 6: N° de NBI según grupos

Al considerar cada uno de los indicadores de las NBI por separado, se observara que el grupo “bajo” presenta carencias muy superiores a los demás, sobre todo los indicadores espacio para cocinar, acceso a agua potable, educación y artículos de confort (véase cuadro 4). Por otro lado, los grupos considerados “medio” y “alto” no presentan grandes diferencias en la mayoría de los indicadores. La excepción a esto se encuentra en los primeros años en los indicadores espacio para cocinar y artículos de confort. A partir de este análisis se puede ver que, en cuanto a los artículos de confort, la partición creada genera una división casi binaria, donde se genera un grupo con carencias importantes en casi todos los indicadores, y otro un grupo (con poca heterogeneidad) con carencias no muy elevadas. En cuanto a la variación temporal, se puede ver que el grupo “bajo” evoluciona negativamente en los indicadores de educación (12,5 % en 2008 y 18,6 % en 2012), servi-

cios higiénicos (12,6 % en 2008 y 18,2 % en 2012), espacio para cocinar (30,7 % en 2008 y 44,1 % en 2012) y espacio habitable (15,5 % en 2008 y 29,5 % en 2012). Como contraparte, se puede ver que el grupo “medio” presenta un leve descenso en los indicadores correspondientes a agua potable (5,3 % en 2008 y 1,7 % en 2012) y artículo de confort (10,9 % en 2008 y 7,5 % en 2012) .

En última instancia se analizó, en cada uno de los grupos, algunos de los indicadores ya presentados inicialmente como descripción de los hogares: sexo y ascendencia del jefe del hogar, número de niños menores de 6 años y tipo de hogar.

Al observar la tabla 5 se puede ver que en cuanto al sexo del jefe del hogar, que el grupo catalogado como “medio” presenta, a lo largo del período, una distribución mas equitativa que los otros dos grupos, donde se nota un predominio del género masculino. Al examinar la ascendencia se puede ver que, en los grupo “medio” y “alto”, el jefe se considera de raza blanca en el 95 % de los casos aproximadamente, mientras que en el grupo “bajo” esta cifra desciende al 85 %, donde el 10 % restante pasa a la categoría de afrodescendientes. En estas variables no se distingue una evolución temporal como en las variables de confort.<sup>9</sup>

Al examinar el número de niños por hogar (véase cuadro 6) se observa una clara distinción en cuanto a los hogares correspondientes al grupo “medio” en tanto que, más del 90 % de los hogares que componen este grupo, no tienen niños menores de 6 años. Por otro lado, se da un contrapunto entre los hogares de los grupos “bajo” y “alto”. En el caso de los hogares del grupo “bajo” puede verse que casi el 10 % de los hogares que lo componen tienen 2 o más niños mientras que, en los hogares del grupo alto, esta cifra sólo asciende al 4 % aproximadamente.

Por último se exploró la composición de los grupos según la tipología de los hogares que los componen. la tabla 7 presenta la distribución de las tipologías en cada grupo para cada año del período de análisis.

Al observar este cuadro, y compararlo con el anterior, surge la explicación de por qué los hogares del grupo “medio” no presentan niños. Esto se debe a que uno de cada tres hogares de este grupo es unipersonal, cifra que es casi el doble de la correspondiente a los otros dos grupos. Por otro lado se ve una diferencia de alrededor de 10 puntos en cuanto los hogares completos del grupo “alto” con respecto al “bajo”. Esta diferencia parece explicarse debido a que este último grupo presenta mayor cantidad de hogares

Tabla 4: Indicadores de NBI por grupo para cada año

	2008	2009	2010	2011	2012
Materialidad					
Bajo	3,6 %	7,1 %	5,2 %	6,9 %	5,5 %
Medio	0,1 %	0,7 %	0,1 %	0,3 %	0,0 %
Alto	0,0 %	0,1 %	0,1 %	0,0 %	0,1 %
Espacio habitable					
Bajo	15,5 %	23,2 %	19,5 %	22,9 %	29,5 %
Medio	0,9 %	2,0 %	0,6 %	1,2 %	1,1 %
Alto	0,8 %	2,3 %	1,5 %	1,5 %	2,2 %
Espacio cocinar					
Bajo	30,7 %	46,0 %	39,8 %	45,6 %	44,1 %
Medio	5,0 %	8,8 %	4,7 %	6,0 %	3,5 %
Alto	1,8 %	4,3 %	3,4 %	3,5 %	4,5 %
Agua potable					
Bajo	26,2 %	30,4 %	23,6 %	27,8 %	2,1 %
Medio	5,3 %	5,5 %	2,5 %	3,4 %	1,7 %
Alto	5,4 %	3,1 %	2,2 %	2,5 %	2,8 %
Servicios higiénicos					
Bajo	12,6 %	17,4 %	14,1 %	19,0 %	18,2 %
Medio	3,0 %	4,2 %	2,7 %	3,3 %	2,7 %
Alto	2,4 %	2,9 %	2,6 %	3,1 %	3,0 %
Energía eléctrica					
Bajo	4,4 %	6,6 %	4,7 %	6,7 %	3,2 %
Medio	0,4 %	1,0 %	0,3 %	0,6 %	0,2 %
Alto	0,3 %	0,6 %	0,4 %	0,4 %	0,5 %
Artículos de confort					
Bajo	46,7 %	63,8 %	53,8 %	60,6 %	55,1 %
Medio	10,9 %	16,2 %	10,4 %	11,7 %	7,5 %
Alto	5,4 %	10,3 %	7,8 %	7,2 %	8,6 %
Educación					
Bajo	12,5 %	13,9 %	13,3 %	15,9 %	18,6 %
Medio	2,0 %	2,5 %	1,7 %	2,6 %	2,4 %
Alto	3,4 %	4,3 %	4,3 %	4,9 %	4,4 %

Tabla 5: Sexo y ascendencia del jefe del hogar

Año	Grupo	Sexo del jefe del hogar		Ascendencia del jefe del hogar		
		Femenino	Masculino	Otra	Afro	Blanco
2008	Bajo	34,8 %	65,2 %	0,3 %	14,6 %	85,1 %
	Medio	51,2 %	48,8 %	0,1 %	4,9 %	95,0 %
	Alto	29,8 %	70,2 %	0,2 %	6,2 %	93,6 %
2009	Bajo	33,8 %	66,2 %	0,3 %	13,1 %	86,5 %
	Medio	53,5 %	46,5 %	0,2 %	4,8 %	95,1 %
	Alto	30,2 %	69,8 %	0,2 %	5,3 %	94,5 %
2010	bajo	34,5 %	65,5 %	0,3 %	13,3 %	86,3 %
	Medio	51,5 %	48,5 %	0,2 %	4,7 %	95,1 %
	Alto	32,1 %	67,9 %	0,1 %	5,7 %	94,2 %
2011	Bajo	37,1 %	62,9 %	0,3 %	13,1 %	86,6 %
	Medio	48,5 %	51,5 %	0,2 %	5,2 %	94,6 %
	Alto	35,4 %	64,6 %	0,2 %	5,6 %	94,2 %
2012	Bajo	38,5 %	61,5 %	0,4 %	15,0 %	84,6 %
	Medio	50,6 %	49,4 %	0,2 %	5,6 %	94,2 %
	Alto	37,5 %	62,5 %	0,3 %	6,6 %	93,2 %

Tabla 6: N° de niños según grupo por año

Año	Grupo	Cero	Uno	Dos	Tres o más
2008	Bajo	70,5 %	20,1 %	7,1 %	2,4 %
	Medio	94,8 %	4,0 %	0,9 %	0,2 %
	Alto	78,1 %	17,4 %	4,1 %	0,4 %
2009	Bajo	71,4 %	19,5 %	7,0 %	2,2 %
	Medio	95,3 %	3,7 %	0,8 %	0,1 %
	Alto	77,7 %	17,6 %	4,2 %	0,5 %
2010	Bajo	71,0 %	20,0 %	6,6 %	2,3 %
	Medio	95,5 %	3,4 %	0,9 %	0,2 %
	Alto	78,3 %	17,4 %	4,0 %	0,4 %
2011	Bajo	71,0 %	20,0 %	6,6 %	1,9 %
	Medio	95,5 %	3,4 %	0,9 %	0,3 %
	Alto	78,3 %	17,4 %	4,0 %	0,3 %
2012	Bajo	69,6 %	20,6 %	7,6 %	2,1 %
	Medio	92,7 %	5,5 %	1,4 %	0,3 %
	Alto	79,1 %	16,5 %	4,0 %	0,3 %

Tabla 7: Tipo de hogar según grupo por año

año	grupo	unipersonal	completo	monoparental	extendido	compuesto
2008	bajo	16,6 %	51,7 %	13,7 %	15,6 %	2,4 %
	medio	33,7 %	36,4 %	10,9 %	16,8 %	2,1 %
	alto	15,0 %	58,7 %	11,6 %	12,9 %	1,8 %
2009	bajo	19,0 %	50,6 %	13,4 %	14,8 %	2,1 %
	medio	35,2 %	34,7 %	12,1 %	15,9 %	2,1 %
	alto	15,1 %	58,9 %	11,4 %	12,8 %	1,8 %
2010	bajo	17,8 %	48,9 %	14,3 %	16,5 %	2,6 %
	medio	36,3 %	36,2 %	10,8 %	14,9 %	1,8 %
	alto	15,5 %	58,5 %	11,8 %	12,5 %	1,7 %
2011	bajo	17,9 %	47,3 %	15,3 %	16,7 %	2,7 %
	medio	31,7 %	36,6 %	12,5 %	17,2 %	1,9 %
	alto	16,1 %	57,7 %	11,5 %	12,9 %	1,8 %
2012	bajo	15,3 %	45,2 %	17,0 %	19,7 %	2,8 %
	medio	30,9 %	38,1 %	11,6 %	17,5 %	1,9 %
	alto	15,4 %	57,2 %	12,1 %	13,4 %	1,9 %

monoparentales y extendidos.

## 5. Conclusiones

Al aplicar la metodología de análisis de cluster se logró identificar tres tipologías de hogares dentro de la muestra de la ECH. La metodología aquí empleada no sólo permite asignar hogares a uno u otro grupo (considerando la máxima probabilidad a posteriori), sino que además, ofrece un “grado de pertenencia” a cada uno de los grupos. Los grupos identificados permiten construir un gradiente de vulnerabilidad. El más vulnerable (denominado “bajo”) se compone por aproximadamente el 30 % de los hogares y corresponde a una situación de bajos ingresos, escolaridad baja y presencia de una o más NBIs. Por otro lado el grupo identificado como “medio” se compone mayoritariamente por hogares sin niños y, a diferencia del grupo “bajo”, presenta mejores condiciones de ingresos/clima educativo.

En última instancia, el grupo denominado “alto” se compone por el 45 % de los hogares que se encuentran en una situación más favorable en cuanto al ingreso per cápita y al clima educativo. Casi el 60 % de los hogares se encuentran en el cuarto o quinto quintil de ingreso y el número de privaciones (NBI) es muy reducido (menos del 10 % de los hogares poseen una o más NBI).

## Referencias

- Akaike, H. (1974). A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*, 19(6):716–723.
- Alkire, S. U. d. O. y Foster, J. U. d. V. (2007). Recuento y medición multidimensional de la pobreza.
- Calvo, J. J., Victor, B., Wanda, C., Paula, C., de los Campos, H., Koolhas, M., Macadar, D., Nathan, M., Nuñez, S., Pardo, I., Tenenbaum, M., y Varela, C. (2011). *Atlas sociodemográfico y de la desigualdad del Uruguay*. TRILCE.
- CONEVAL (2010). Metodología para la medición multidimensional de la pobreza en México. Technical report, México DF.
- Dempster, A. P., Laird, N. M., y Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *JOURNAL OF THE ROYAL STATISTICAL SOCIETY, SERIES B*, 39(1):1–38.
- INE (2013). Estimación de la pobreza por el método del ingreso. Technical report, Instituto Nacional de Estadística, Montevideo, Uruguay.
- Katzman, R. (1989). Revista de la. *Revista de la CEPAL*, (37):141–152.
- Klugman, J. P. (2010). *Informe sobre Desarrollo Humano 2010 La verdadera riqueza de las naciones : Caminos al desarrollo humano*. Mundi-pren edición.
- Moustaki, I. y Papageorgiou, I. (2005). Latent class models for mixed variables with applications in Archaeometry. *Computational Statistics & Data Analysis*, 48(3):659–675.
- R Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464.

Wolff, E. N., Zacharias, A., Masterson, T., y Eren, S. (2012). A comparison of Inequality and living standards in Canada and the United States using Eepanded measure of economic well-being.



# Instituto de Estadística

---

## Serie Documentos de Trabajo



FACULTAD DE  
CIENCIAS ECONÓMICAS  
Y DE ADMINISTRACIÓN

**IESTA 80** ANIVERSARIO INSTITUTO  
DE ESTADÍSTICA



UNIVERSIDAD  
DE LA REPÚBLICA  
URUGUAY

Gonzalo Ramirez 1926, Piso 1, Oficina 23 - C.P. 11200 -  
Montevideo, Uruguay  
Teléfono: (598) 2410 2564  
<https://iesta.fcea.udelar.edu.uy/>  
Área Publicaciones

1er semestre, 2022

Nº 2/22