
Una aplicación del análisis factorial múltiple para el estudio de la pobreza multidimensional en Uruguay

Maximiliano Saldaña

Laura Nalbarte

Ramón Álvarez-Vaz

Serie Documentos de Trabajo

Nº4/22

2022

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 - NoCo-
mercial - CompartirIgual 4.0 Internacional.

Forma de citación sugerida para este documento:

Maximiliano Saldaña, Laura Nalbarte, Ramón Álvarez-Vaz (2022). *Una Aplicación del análisis factorial múltiple para el estudio de la Pobreza multidimensional en Uruguay* (Serie Documentos de Trabajo; Nº4/22). 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>

Una Aplicación del análisis factorial múltiple para el estudio de la Pobreza multidimensional en Uruguay

Maximiliano Saldaña ¹; Laura Nalbarte ² Ramón Alvarez-Vaz ³;

Departamento de Métodos Cuantitativos, Instituto de Estadística,

Facultad de Ciencias Económicas y de Administración,

Universidad de la República

Resumen

La medición de la pobreza multidimensional ha sido un tema de interés tanto en el marco internacional como el nacional. En el presente trabajo se explora el Análisis Factorial Múltiple -AFM- (Escofier y Pages, 1994) como una opción para resumir la información relativa a las distintas dimensiones de la pobreza. Se busca llegar a una propuesta de un índice de pobreza multidimensional basado en el AFM como una alternativa al empleado por la Organización de las Naciones Unidas. La metodología que se aplica es considerada particularmente pertinente para el objetivo dada su capacidad de resumir información de conjuntos extensos de variables, que son ponderados apropiadamente para obtener como resultado final un conjunto menor de nuevas variables, con la menor pérdida de variabilidad explicada posible. A su vez, esta metodología permite visualizar vínculos entre las variables y entre las distintas dimensiones de la pobreza. Se hace uso de bases de datos de la Encuesta de Gastos e Ingresos de los Hogares 2016-2017, realizada por el Instituto Nacional de Estadística en Uruguay.

Palabras clave: Pobreza Multidimensional, Análisis Factorial Múltiple, Índice de Pobreza ;

CÓDIGOS JEL: C10, C18, C38, C83 .

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

¹ *email:* maximiliano.saldana@fcea.edu.uy, ORCID: 0000-0002-9137-3579

² *email:* laura.nalbarte@fcea.edu.uy, ORCID:

³ *email:* ramon.alvarez@fcea.edu.uy, ORCID: 0000-0002-2505-4238

ABSTRACT

The measurement of multidimensional poverty has been a topic of interest both in the international and national framework. In this paper, Multiple Factor Analysis -AFM- (Escofier and Pages, 1994) is explored as an option for summarize the information related to the different dimensions of poverty. It seeks to arrive at a proposal for a multidimensional poverty index based on the AFM as an alternative to the one used by the United Nations Organization. The applied methodology is considered particularly pertinent to the objective given its ability to summarize information from extensive sets of variables, which are appropriately weighted to obtain as a final result a smaller set of new variables, with the least possible loss of explained variability. In turn, this methodology makes it possible to visualize links between the variables and between the different dimensions of poverty. Use is made of databases from the Household Expenditure and Income Survey 2016-2017, carried out by the National Institute of Statistics in Uruguay.

Key words:Multidimensional Poverty, Multiple Factor Analysis, Poverty Index .

JEL CODES:C10, C18, C38, C83 .

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

1. Introducción

La medición de la pobreza representa un tema de interés para cualquier sociedad, constituye una manera de constatar el nivel de bienestar de sus integrantes y para contar con información para la toma de acciones en pos de su beneficio.

Las metodologías para efectuar dicha medición son variadas y encaran el problema desde distintos enfoques, en particular, se destaca la distinción de las medidas de corte unidimensional y las llamadas multidimensionales. Las primeras buscan identificar el estatus de pobreza de una persona a partir de una sola variable, siendo uno de los métodos más comunes el del ingreso, en el que se calculan los ingresos que necesitaría un hogar para satisfacer las necesidades de sus integrantes y se comparan con sus ingresos efectivos (1). Estas medidas son aplicables de manera relativamente sencilla e indudablemente son de importancia, permitiendo tener una noción general del bienestar del individuo, en el caso del método del ingreso dándole todo el peso sobre este bienestar al dinero percibido por el hogar. No obstante, ha resultado de interés el estudio de medidas multidimensionales las cuales permitan tener en cuenta un conjunto de variables consideradas de incidencia en la pobreza, tratando de capturar sus distintas aristas y de tener una visión más amplia de ella. Desde el año 2010 a nivel internacional se emplea el Índice de Pobreza Multidimensional (2) con este mismo fin, que resume la información de distintas dimensiones de la pobreza haciendo uso de un promedio ponderado de variables consideradas representativas de cada una de ellas.

En el presente trabajo se presenta un estudio multidimensional de la pobreza en Uruguay haciendo uso de información recabada en la Encuesta Nacional de Gastos e Ingresos de los Hogares 2016-2017 y empleando la metodología del análisis factorial. Se busca llegar a una propuesta de un índice de pobreza multidimensional alternativo basado en el análisis factorial múltiple -AFM- (3), metodología considerada particularmente apropiada para el objetivo dada su capacidad de resumir información de conjuntos de variables. El AFM posee la virtud de permitir la inclusión de un conjunto extenso de variables y obtener como resultado de la técnica un conjunto menor de nuevas variables, con la menor pérdida de información posible. A su vez, esta metodología permite visualizar vínculos

entre las variables y las distintas dimensiones de la pobreza.

Con respecto al ciclo de vida de este trabajo se deja constancia que un primer avance de esta investigación fue presentado en el Clatse 2021 en Montevideo y luego en junio de 2022 se presentó un resumen extendido para la XV Semana Internacional de la Estadística y la Probabilidad, de la Facultad de Ciencias Físico Matemáticas de la Benemérita Universidad Autónoma de Puebla con número de preprint [10.31219/osf.io/7sqvc](https://doi.org/10.31219/osf.io/7sqvc); se complementa posteriormente con el trabajo casi terminado previo a la elaboración de este DT, el que se presentó en las VI Jornadas Argentinas de Econometría de la FCE de la UBA, Buenos Aires, Argentina, en setiembre 2022 que aparece en <https://doi.org/10.5281/zenodo.7145264>. El código computacional y los datos en los que se basa este trabajo está disponible en el repositorio <https://gitlab.com/iesta.fcea.udelar/estudio-multidimensional-de-la-pobreza-para-engih2016-2017>.

La organización del documento es la siguiente: en la primera sección se introduce el problema bajo estudio, para luego presentar una reseña de los antecedentes; en la tercera se presenta brevemente la metodología empleada, en la cuarta las aplicaciones y resultados de los diversos análisis factoriales para finalmente presentarse las conclusiones en la quinta sección.

2. Antecedentes

En el ámbito internacional la medida de pobreza multidimensional estandar es la empleada por el Programa de las Naciones Unidas en sus Reportes de Desarrollo Humano el Índice de Pobreza Multidimensional (MPI) desarrollado en (2) y siendo su última versión la presentada en (4). Este índice hace uso de un puntaje de privación ponderado, el cual se obtiene a partir de un promedio ponderado de tres conjuntos de variables indicadoras. La unidad de análisis es el hogar que será identificado como “multidimensionalmente pobre” si su puntaje de privación supera un punto de corte. La medida final toma en cuenta tanto la proporción de hogares considerados pobres como el nivel promedio de la pobreza, de manera de captar la intensidad de la pobreza. Este índice fue creado con la intención de crear un resumen de la situación de pobreza de una región, el cual puede ser

descompuesto por dimensiones y puede ser comparado con la medida de otras regiones y de otros momentos en el tiempo.

La metodología del Análisis Factorial ha sido aplicada previamente al estudio de la pobreza multidimensional, tanto para la creación de índices como para obtener insumos para otras metodologías.

En (5) se realiza un estudio multidimensional de la pobreza para países de Medio Oriente y África del Norte, mediante un Análisis de Componentes Principales (ACP) y Análisis de Cluster, con la intención de definir perfiles de pobreza de los distintos países de la región. Se considera de un conjunto de variables cuantitativas referentes a distintas dimensiones de pobreza, resumiendo su información en tres factores del análisis factorial realizado para posteriormente emplearlos para separar los países en grupos mediante el análisis de cluster.

Por otro lado en (6) se busca construir un índice sintético de pobreza empleando datos de Venezuela. Mediante el Análisis de Homogeneidad se obtienen variables cuantitativas a partir de categóricas, con las que se hacen distintos ACP con la finalidad de obtener índices parciales, tomados como las primeras componentes de los análisis, los cuales luego son agregados en una suma ponderada para obtener el índice de pobreza final. Esta metodología cuenta con el problema que no es tomado en cuenta que las variables que en última instancia son empleadas en el índice provienen de análisis distintos y como tales deben ser ponderados apropiadamente.

En el Uruguay hay dos medidas principales de la pobreza, ambas realizadas por el Instituto Nacional de Estadística (INE): el método del ingreso y el de las necesidades básicas insatisfechas (NBI). La primera de ellas, descrita en (1), hace uso de una canasta básica alimentaria y otra no alimentaria definidas a partir de la Encuesta Continua de Hogares del INE, con la finalidad de definir los ingresos que un hogar requiere para satisfacer sus necesidades. Si el hogar no cuenta con los ingresos suficientes, teniéndose en cuenta la cantidad de personas del hogar, entonces se le considerará pobre y si cuenta con ellos no pobre. La segunda metodología es de carácter multidimensional y “Se orienta a identificar la falta de acceso a bienes y servicios (o problemas críticos en cuanto a la calidad de

aquellos) cuya disposición constituye una condición para el ejercicio de derechos sociales” (7). Toma en cuenta seis dimensiones en las cuales se definen variables indicadoras de necesidades insatisfechas y finalmente se considera la cantidad de ellas en el hogar.

Un acercamiento al problema de la medición de la pobreza multidimensional para el caso de Uruguay se trabaja en (8). Se emplean datos de la Encuesta Nacional de Hogares Ampliada (ENHA) 2006 del INE y la metodología es similar a la empleada en (5), pero haciendo uso de conjuntos de variables cualitativas los cuales son resumidos en variables cuantitativas mediante Análisis de Correspondencias Múltiples, que posteriormente son empleadas en un Análisis de Cluster para identificar un grupo de hogares de pobreza multidimensional.

Finalmente, en (9) se trabaja en sintetizar la situación de pobreza de los hogares en un índice de pobreza multidimensional empleando el análisis factorial y en construir tipologías de hogares a partir del análisis de cluster, haciendo uso de datos de la Encuesta Continua de Hogares (ECH) del INE de los años 2008 a 2012. A partir de cuatro ACM se obtienen variables que resumen conjuntos de variables que representan distintas dimensiones de pobreza, para luego emplearlas en un ACP, ponderando las variables de resumen haciendo uso de la inercia del ACM del cual provenían. El problema de esta forma de ponderación, como se plantea en (10), es que no se estaría tomando en cuenta el enfoque del análisis factorial, en donde se buscan las direcciones de máxima inercia. Si se ponderan los primeros factores de cada ACM por el inverso de la inercia total de su correspondiente análisis no se toma en cuenta que dicha inercia puede estar más o menos distribuída en los distintos factores de cada análisis. Para solucionar esto (3) proponen un sistema de ponderación distinto enmarcado en el análisis factorial múltiple que plantean, donde la ponderación se hace con la inercia de los primeros ejes de cada grupo de variables. Esta metodología se desarrollará y aplicará en el presente documento.

3. Metodología

En (12) se presenta un desarrollo sobre la importancia de la fase de selección de dimensiones al medir la pobreza multidimensional, desde el enfoque de las capacidades de Amartya

Sen, donde la pobreza es definida como una privación de libertades para poder realizar sus funcionamientos (las cosas que quieren ser y hacer). Desde este punto de vista de carácter normativo se identifican cinco métodos de selección empleados por los investigadores:

1. Basado en los datos existentes o en alguna convención.
2. A partir de supuestos de lo que las personas valoran o deberían valorar. “Conjeturas informadas” de los investigadores.
3. Tomando en cuenta un consenso público.
4. Realizar un proceso deliberativo participativo y continuo, cada tanto actualizándose.
5. Emplear evidencia empírica para definir lo que las personas valoran, por ejemplo, analizando el consumo de las personas.

Alkire enfatiza que estas metodologías no están aisladas y en general tienen interacción entre ellas, empleándose varias de ellas al mismo tiempo. A su vez, la autora plantea la importancia de permitir cierta flexibilidad a la hora de generar dimensiones, dado que distintos contextos tendrán necesidades distintas, más allá del atractivo de contar con un listado “canónico” fijo que permita la comparación en distintos lugares y tiempos.

Tomando en cuenta la disponibilidad de los datos y un enfoque normativo de las dimensiones de importancia en el bienestar de las personas, para seleccionar las variables y agruparlas en dimensiones se consideraron como base aquellas presentadas en (9), añadiéndose nuevas variables consideradas de relevancia. Las dimensiones que se trabajan son: tecnologías de información y comunicación (TICs), confort, vivienda y laboral-educativa. No obstante, quedan abiertas a discusión para futuras aplicaciones.

Adicionalmente, se consideraron algunas variables fuera de estos conjuntos para observar la distribución del índice generado en sus grupos, estas son: pobreza subjetiva, pobreza por método del ingreso y sexo del jefe/a del hogar.

3.1. Análisis Factorial

Las distintas técnicas de análisis factorial, como lo son el análisis de componentes principales, análisis de correspondencias simples, análisis de correspondencias múltiples, análisis factorial múltiple, etc., parten de un mismo conjunto de nociones.

Se toman las nubes de puntos que representan a las filas y columnas de la matriz de datos a analizar y se proyectan en dos nuevos espacios óptimos de menor dimensión, siendo el criterio de optimalidad que cumplan la condición de maximizar la inercia explicada (que se preserve la mayor proporción de la variabilidad original de los datos). Luego, las nubes originales son representadas en los nuevos espacios obtenidos, estando ambas representaciones vinculadas (16).

Como resultado se obtienen ejes de inercia, nuevas variables que resultan de la combinación lineal de las originales y que son ortogonales entre sí. Este tipo de técnicas permiten resumir la información esencial de los datos originales, eliminando la redundancia y pasándose a un conjunto menor de dimensiones a estudiar.

Dada una matriz de datos del tipo individuos \times variables, tanto los primeros como las segundas pueden ser ponderados, por ejemplo para tomar en cuenta que los datos provienen de un diseño muestral, como el caso de el presente trabajo, o que las variables tienen distinto peso, lo cual también será aplicado.

3.2. Análisis Factorial Múltiple

El Análisis Factorial Múltiple (AFM) fue desarrollado en la década del noventa por Brigitte Escofier y Jérôme Pàges, en particular en (3) se encuentra explicada la metodología junto a un ejemplo de aplicación.

La técnica es un desarrollo del análisis factorial, donde el elemento clave de los métodos es que las variables se colocan en grupos y pesos son aplicados a cada uno de ellos, para poder hacer comparable la influencia de cada grupo en el análisis global.

3.2.1. AFM con variables cualitativas

Dada la naturaleza cualitativa de las variables consideradas para el análisis se desarrolla la metodología del AFM para este tipo de variables.

Se considera el enfoque que considera que los datos se tienen en una Tabla Disjunta Completa (TDC), es decir, una matriz de ceros y unos donde las filas corresponden a los individuos y las columnas a las modalidades de las variables. Para la modalidad de cada variable que el individuo tiene el valor de la matriz es uno, el resto de las modalidades tendrán valor cero. El término general de la TDC se representa mediante y_{ik} .

En el AFM de variables cualitativas ahora las variables indicadoras se consideran en conjunto en grupos que incluyen múltiples variables. No se hace distinción entre grupos de indicadores y grupos de variables bajo este enfoque (10).

Al igual que en el caso de que se tuvieran variables cuantitativas, resulta necesario balancear la influencia de los grupos de variables en un análisis general, teniendo en cuenta que un grupo de mayor dimensión tendrá una influencia más dispersa sobre más ejes que un grupo más reducido. Esto se logra empleando las inercias de las direcciones principales de inercia de los distintos grupos en la ponderación, llevándose la inercia máxima de las subnubes asociadas a un grupo de variables a 1.

El AFM para variables cualitativas sigue la lógica del ACM considerando ponderaciones de las categorías de las variables. Para hacer que la inercia máxima de la nube asociada al grupo j de variables sea 1, dividimos el peso de cada categoría de dicho grupo entre λ_1^j (la inercia del primer eje de un ACM que considera solo al grupo j). Desde la perspectiva de la nube de individuos, entonces la distancia entre dos individuos i^j y l^j , considerándose el grupo de variables j es:

$$d^2(i^j, l^j) = \frac{1}{Q_j \lambda_1^j} \sum_{k \in K_j} p_k \left(\frac{y_{ik}}{p_k} - \frac{y_{lk}}{p_k} \right)^2 \quad (1)$$

Donde Q_j es el número de variables del grupo j , K_j es el número de categorías del grupo

j y $p_k = \sum_i p_i y_{ik}$ es el peso del conjunto de individuos que poseen la categoría k .

Considerándose todos los grupos de variables, las distancias entre los individuos i y l de la nube de individuos resulta:

$$d^2(i, l) = \sum_j d^2(i^j, l^j) = \sum_j \frac{1}{Q_j \lambda_1^j} \sum_{k \in K_j} p_k \left(\frac{y_{ik}}{p_k} - \frac{y_{lk}}{p_k} \right)^2 \quad (2)$$

$$= \sum_j \sum_{k \in K_j} \frac{1}{Q_j \lambda_1^j} p_k \left(\frac{y_{ik}}{p_k} - \frac{y_{lk}}{p_k} \right)^2. \quad (3)$$

En definitiva, el AFM de variables cualitativos agrupados es un ACM donde el peso de las variables del grupo j es $1/(Q_j \lambda_1^j)$.

Al igual que en cualquier tipo de análisis factorial se pueden pasar de las coordenadas en el eje de inercia s de los individuos ($F_s(i)$) a las coordenadas de las variables (en este caso de las categorías de las mismas) en dicho eje ($G_s(k)$), mediante las fórmulas de transición. En particular, una categoría k se encuentra a menos de un coeficiente en el baricentro de los individuos que la poseen:

$$G_s(k) = \frac{1}{\sqrt{\lambda_s}} \frac{1}{p_k} \sum_i p_i y_{ik} F_s(i). \quad (4)$$

También se puede obtener la fórmula de transición en el otro sentido. Un individuo se encuentra a menos de un coeficiente en el baricentro de las categorías que posee:

$$F_s(i) = \frac{1}{\lambda_s} \sum_j \frac{1}{Q_j \lambda_1^j} \sum_{k \in K_j} y_{ik} G_s(k). \quad (5)$$

3.2.2. Forma alternativa de ver el AFM

Los resultados finales que se obtienen al realizar un AFM sobre un conjunto de variables agrupadas resultan equivalentes a considerar un análisis de componentes principales empleando los factores obtenidos en los distintos análisis iniciales por grupo de variables, ponderando apropiadamente dichos factores. La ponderación del factor s del grupo j de

variables, en el caso que se estandaricen los factores, es el cociente de inercias λ_s^j/λ_1^j . El numerador es la inercia del eje a ponderar mientras que el denominador es la inercia del primer eje (la dirección principal de inercia).

Bajo esta presentación alternativa del AFM los factores de los análisis iniciales son variables activas del análisis y su inercia proyectada puede ser interpretada como su contribución a la inercia de los ejes del análisis final.

3.3. Aplicación

Los datos utilizados provienen de la Encuesta Nacional de Gastos e Ingresos de los hogares (2016-2017), efectuada en el Uruguay por el INE. En dicha encuesta se relevó información sobre el consumo e ingreso de los hogares de todo el país, junto a información referente a sus características socio-demográficas, materiales y de autopercepción de la pobreza. Se cuenta con datos muestrales a nivel hogar y nivel persona, que surgen de un diseño muestral complejo, donde se relevó la información de 6889 hogares del territorio nacional (11). Se opta por trabajar a nivel hogar, resumiendo variables consideradas pertinentes que se preguntaban a nivel persona.

Para el procesamiento de los datos fue empleado el software R (13), haciéndose uso del conjunto de paquetes *Tidyverse* (14) y para los análisis factoriales se hizo uso del paquete *FactoMineR* (15).

4. Resultados

4.1. Análisis descriptivo inicial

En un principio se realizan análisis de correspondencias múltiples para estudiar el comportamiento y relaciones de las modalidades de las variables y ver la factibilidad de resumir la información de cada dimensión. Si el AFM se encara desde el segundo enfoque presentado en la sección anterior, lo que se busca es obtener un conjunto reducido de variables cuantitativas (una selección de los factores resultantes de los ACM), para posteriormente incluirlas en un análisis de componentes principales, incorporando así en un mismo análisis la información de las distintas dimensiones de la pobreza. Desde esta perspectiva se

estaría realizando un resumen por fases, primero resumiendo por dimensión para posteriormente obtener un resumen de la pobreza en general.

Se trabaja en primera instancia con los primeros planos factoriales resultantes de los distintos ACM considerados, dado el alto porcentaje de inercia que acumulan y la simplicidad de representación e interpretación gráfica. El porcentaje de inercia explicada acumulada por las primeras dos dimensiones no fue menor a 72% una vez realizado el ajuste de Benzecri, una corrección que se hace para obtener una mejor aproximación de la inercia acumulada por cada eje (16). De hecho, en la mayoría de los análisis incluso fue mayor al 98%, lo que implica que los primeros dos ejes retienen gran parte de la variabilidad de la nube de puntos original.

Como conclusión general a los cuatro ACM se puede afirmar que el primer eje de cada uno de ellos permite realizar una distinción entre los hogares del país. Aquellos que cuentan con menos carencias en la dimensión considerada tendrán en promedio menores valores del primer eje, mientras aquellos que cuentan con un mayor número de carencias presentarán mayores valores del eje. A su vez, al proyectar como variables suplementarias la pobreza por el método del ingreso y la pobreza subjetiva, las categorías que refieren al estatus de hogar pobre se corresponden con mayores valores del primer eje, aunque dicho vínculo no resulta claro, por ejemplo en el caso de la dimensión empleo y educación.

A modo de ejemplo, si se observa el gráfico correspondiente al ACM de la dimensión confort de la Figura 1, el primer eje permite distinguir entre aquellos hogares que carecen de artefactos de confort como el calefón, refrigerador, lavadora y microondas de aquellos que no, contando los primeros con valores elevados del primer eje y los segundos con valores menores. En el gráfico de la Figura 2 se observa como para la dimensión empleo y educación mayores valores del primer eje se corresponderán con una situación de mayor precariedad laboral y educativa.

En la Figura 3 se aprecia que para la dimensión TICs valores mayores del primer eje se corresponden con la falta de acceso a dispositivos tecnológicos, mientras que en el caso de la dimensión vivienda (Figura 4) valores altos de su primer factor reflejan problemas estructurales del hogar.

igual peso, lo cual desde un punto de vista cualitativo puede que no sea deseado, es decir, se puede desear dar más peso a alguna dimensión intencionalmente. El IPM presentado en (2) presenta la posibilidad de extensión al caso donde los distintos grupos de variables tengan una ponderación distinta, aunque en la práctica se emplean ponderaciones iguales de ellos.

Debe tenerse en cuenta que dada la metodología aplicada para obtener un índice de simple interpretación es que la medida que se crea depende del contexto donde se efectúa; en caso de cambiar el conjunto de hogares la medida puede cambiar. Por esto la medida presentada es de pobreza relativa a toda la población, permite decir que una persona es pobre en un determinado contexto, pero no comparar en general con otros, como sí lo permite el IPM de la ONU.

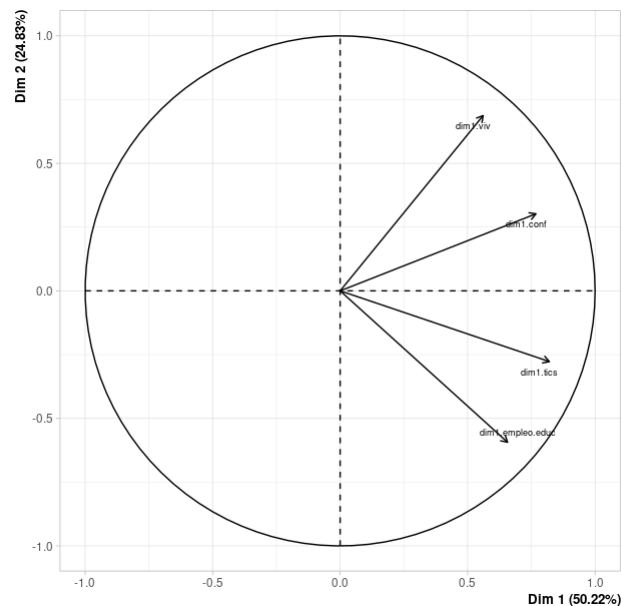


Figura 5: Representación del primer plano factorial, análisis factorial múltiple.

En la Figura 5 se observa la representación gráfica de las proyecciones de los primeros ejes iniciales usados como variables para realizar el AFM, en el primer plano principal de este último. Los dos primeros factores del ACP que serían la segunda fase del AFM acumulan un 75 % de la inercia, donde el primero de ellos acumula el 50 %. Dada la dirección de los vectores que representan las variables en el nuevo subespacio se puede afirmar que

mayores valores en las variables empleadas para resumir las distintas dimensiones de pobreza se verán reflejados en mayores valores del primer eje del AFM. Las coordenadas de la dimensiones, que en el caso presente al ser un ACP pueden ser interpretadas como las correlaciones de las variables iniciales con los factores resultantes son:

- Dimensión TICs: 0.87.
- Dimensión Confort: 0.57.
- Dimensión Vivienda: 0.33.
- Dimensión Empleo y educación: 0.83.

Se puede apreciar que el primer eje se ve mayormente correlacionado con las variables que resumen las dimensiones referentes a las TICs y el Empleo y educación, por lo que puede ser un buen representante de ellas, pero para el resto de las variables la correlación es más débil, siendo su calidad de representación menor. Esta pérdida de información es de esperarse al realizar una reducción de dimensiones, aunque se deberá posteriormente evaluar la validez de realizarla y si resulta necesario tomar otra alternativa que permita retener una mayor cantidad de información.

El índice se genera a partir del primer factor del análisis factorial anterior, que quedó definido de manera tal que un hogar que presente una situación de mayor pobreza multidimensional tenga un valor mayor. A su vez, para facilitar su lectura, el índice se estandariza, restándole su valor mínimo y dividiéndolo entre su máximo, de manera que tome valores entre 0 y 1. El índice obtenido presenta una asimetría positiva, siendo la asimetría no paramétrica de 0.14, su media es de 0.37 y su mediana de 0.34. Se presenta un histograma del IPM en la Figura 6 .

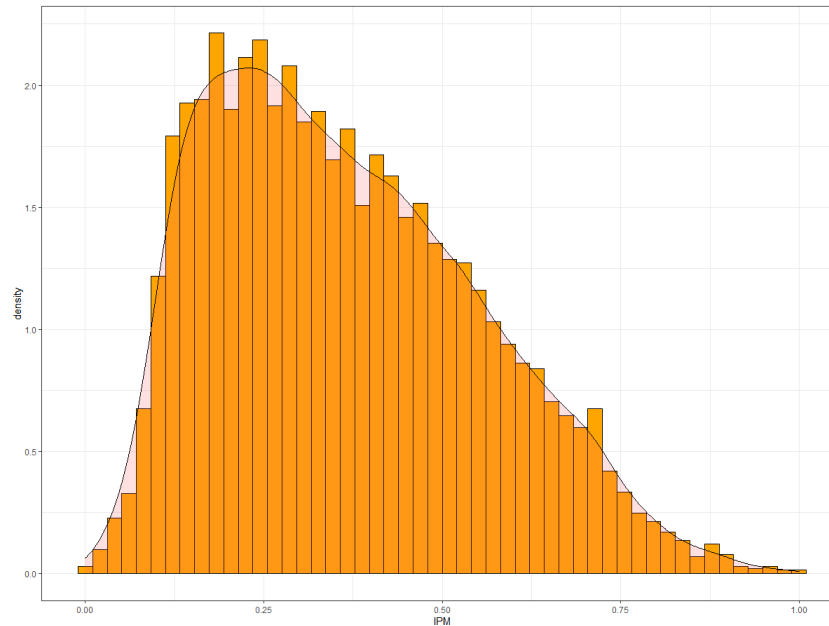


Figura 6: Densidad estimada del índice de pobreza multidimensional.

4.3. Cálculo del índice de pobreza multidimensional (IPM): construcción alternativa

El IPM que se busca crear a través de AFM también puede realizarse desde la primera alternativa planteada en la sección metodología, trabajando a nivel de las variables categóricas iniciales en lo que básicamente es un ACM con ponderaciones de las modalidades. Este caso sería equivalente al análisis anterior si en lugar de los primeros ejes de los análisis iniciales se tomaran en cuenta todos los obtenidos en cada ACM y puede emplearse en el caso en que se desea preservar la mayor cantidad de información en el AFM.

Como puede verse en la Figura 7, en el AFM alternativo cuando se consideran los primeros ejes parciales como variables suplementarias, siguen teniendo una correlación alta con el primer factor del AFM, pero ahora entran en juego las proyecciones de los otros ejes, que presentan una correlación menor con el primer y segundo eje. Además, las direcciones de estas últimas proyecciones no cuenta con una interpretación clara como las consideradas

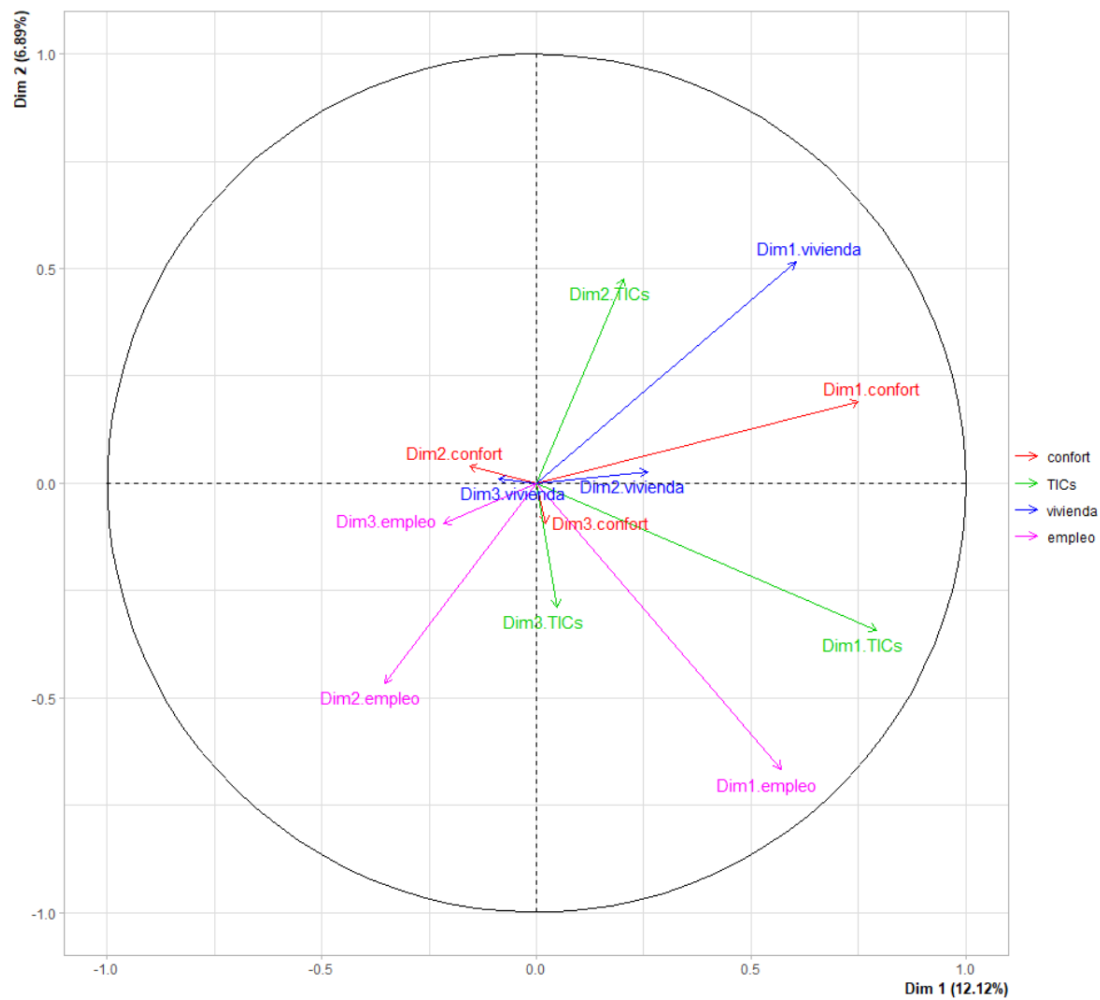


Figura 7: Proyecciones de ejes parciales del AFM alternativo.

inicialmente.

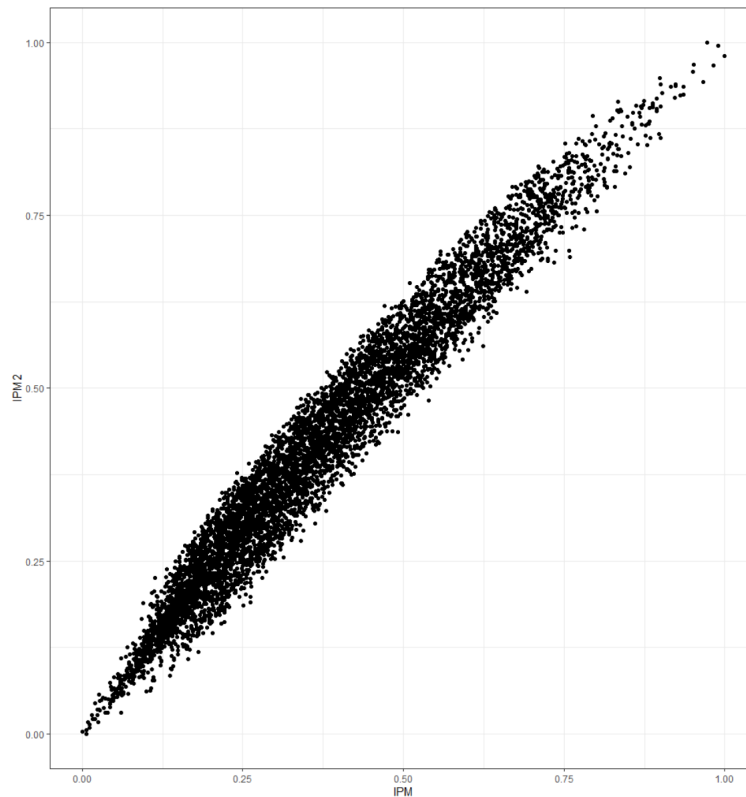


Figura 8: Gráfico de dispersión donde eje X es el IPM de la sección anterior y el eje Y es el alternativo.

Cuando se gráfica el IPM obtenido en primera instancia contra el obtenido mediante la interpretación alternativa y la inclusión de más factores se observa que si bien los resultados de los índices no son idénticos, sí presentan una correlación lineal alta, en particular de 0.98 (Figura 8). Que no sean iguales se debe al aumento en la variabilidad dado por la consideración de un mayor número de variables en la segunda etapa del AFM. Este resultado nos muestra que de alguna forma si bien se pierde información al realizar un AFM como el realizado en primera instancia, de todas maneras el índice representará el ordenamiento de los hogares relativo a la medida de resumen de la pobreza. Por este motivo se opta por trabajar con la primera metodología empleada.

4.4. Discusión: Análisis descriptivo haciendo uso del índice

Una vez que se cuenta con el valor del índice para cada hogar, a modo de ver como se comporta ante distintas caracterizaciones y también de observar si resulta coherente se realiza una serie de descripciones del índice según variables cualitativas. También se estudia el uso del plano para la descripción de la pobreza.

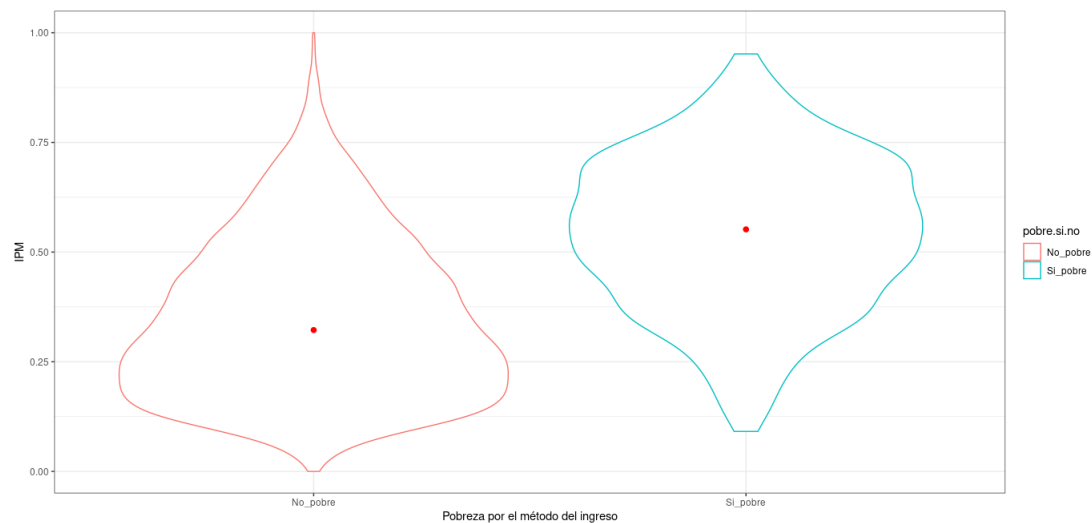


Figura 9: Gráficos de violín del IPM, por status de pobreza usando método del ingreso.

En la Figura 9 se presenta el gráfico de violín del IPM, separando por pobreza por el método del ingreso. Cuando se considera este tipo de pobreza, para el grupo de los clasificados como no pobres la mediana es menor y la distribución se concentra a menores valores que el grupo de los pobres. Esto indica que el índice obtenido tiene cierta correspondencia con el método del ingreso (recordar que mayores valores del índice indican una situación de mayor pobreza multidimensional). El grupo no pobre cuenta con hogares con valores del IPM atípicamente altos, mayores incluso que el máximo del grupo pobre y a su vez tiene una dispersión mayor. Esto puede deberse a que el índice capta otros componentes de la pobreza que no están captados en el ingreso.

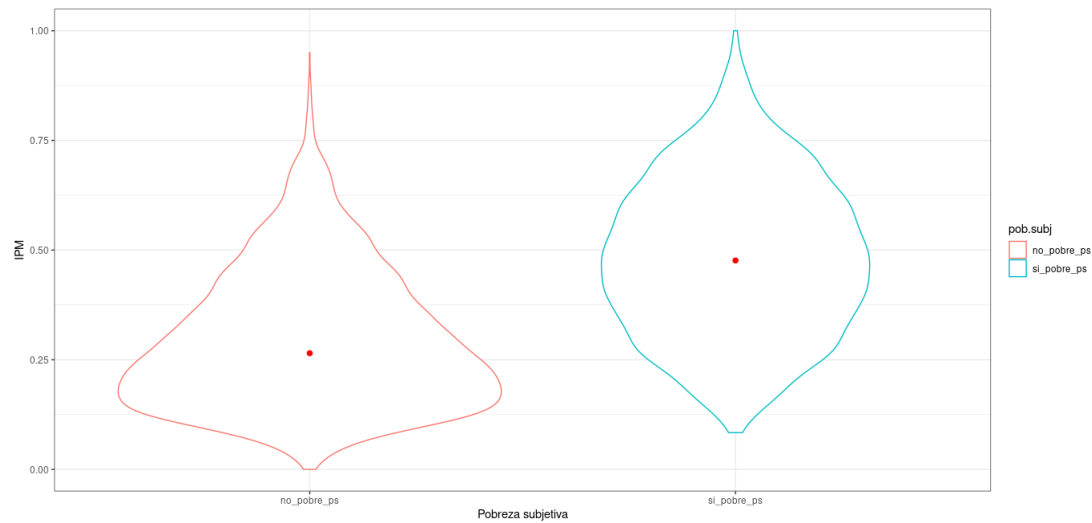


Figura 10: Gráficos de violín del IPM, considerando la pobreza por el método del ingreso.

En la Figura 10, se presenta el mismo gráfico pero ahora para la pobreza subjetiva, la cual a priori se podría pensar que puede reflejar más dimensiones que el ingreso. Los resultados son similares a los anteriormente conseguidos, aunque en este caso el máximo del IPM se encuentra en el grupo de los hogares que se autorreportan pobres y la distribución de los pobres presenta una mayor simetría, mientras que la de los no pobres tiende a acumularse a menores valores.

Retomándose el primer plano factorial (11), a partir de cuyo primer eje se obtuvo el IPM, se puede notar que el primer cuadrante es el que concentra mayor cantidad de hogares pobres por el método del ingreso. A pesar de esto, hay hogares no pobres con valores altos de ambos ejes, lo cual nuevamente se puede deber a que el resumen de las distintas dimensiones de pobreza permite captar carencias que el método del ingreso no toma en cuenta.

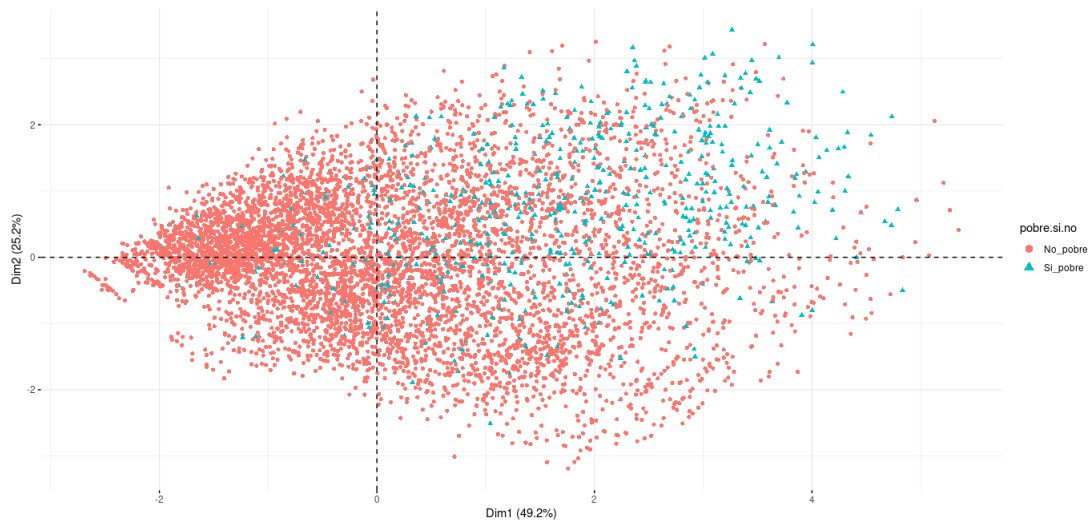


Figura 11: Individuos en primer plano factorial del AFM, coloreados por condición de pobreza (método del ingreso).



Figura 12: Individuos en primer plano factorial del AFM, coloreados por condición de pobreza (método del ingreso).

En cuanto a la pobreza autorreportada, el primer y cuarto cuadrante cuenta con la mayor acumulación de hogares que se consideran pobres, mientras que el segundo y el tercero cuenta con mayor cantidad de hogares que no (Figura 12). Aunque tiene que tenerse en cuenta que la proporción de los hogares que se consideran pobres es mayor a la de los que son pobres por el método del ingreso, los ejes obtenidos cuentan con una mayor correspondencia con este tipo de pobreza, aunque nuevamente no resulta del todo congruente. Este gráfico también pone de manifiesto que puede resultar de interés considerar más de un eje del AFM para la medición de la pobreza.

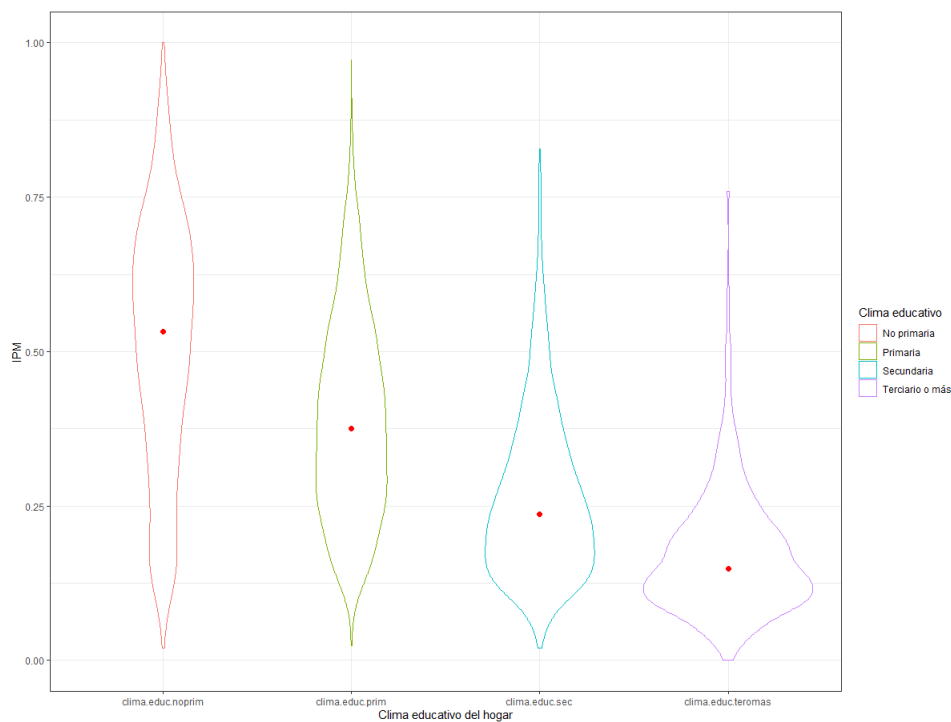


Figura 13: Gráfico de violín del IPM, por clima educativo.

Considerando las distribuciones del IPM para los hogares separando por el clima educativo (la moda del nivel educativo de las personas del hogar), se puede apreciar un ordenamiento claro como muestra la Figura 13 : a mayor nivel educativo la distribución del índice tiende a acumularse a menores valores. Esto resulta coherente dado que esta fue una de

las variables consideradas en la dimensión educación-laboral, pero también muestra su importancia en la pobreza multidimensional.

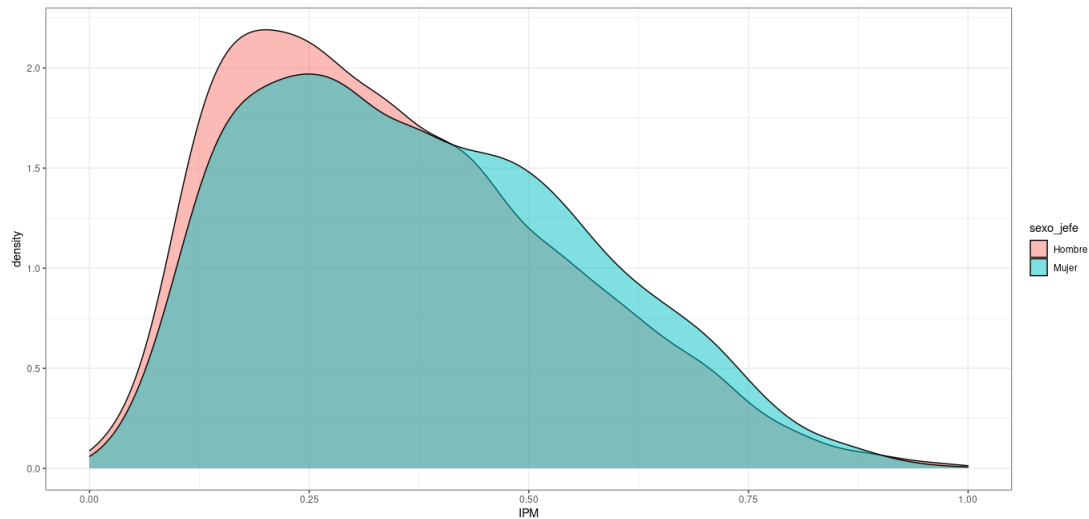


Figura 14: Densidad estimada del IPM, por sexo del jefe del hogar.

En cuanto a la densidad del IPM según el sexo del jefe del hogar, en la Figura 14 se puede notar que las densidades estimadas del IPM para los distintos sexos difieren entre sí. Para los hombres la media resulta menor en comparación a las mujeres (0.35 y 0.38 respectivamente). Al comparar las medianas el resultado es análogo.

5. Conclusiones y consideraciones finales

La metodología del análisis factorial múltiple presenta potencial para la medición de la pobreza, aunque deben ser estudiados los puntos a favor y en contra que presenta. Los distintos análisis iniciales permitieron obtener un resumen de la situación de pobreza de los hogares en cada una de sus respectivas dimensiones, en la forma de sus primeros ejes factoriales. El resto de los ejes no resultaron tan interpretables en primera instancia, pero sí aportan a la hora de estudiar el vínculo entre modalidades de las variables.

El índice obtenido tiene cierta correspondencia con el método del ingreso, aunque capta aspectos distintos al dinero percibido por los hogares. Se pudo observar además que distintas subpoblaciones de la muestra, obtenidas al separar esta última según variables, presentaron distintas características.

Siendo el IPM presentado en (4) el estándar, se considera pertinente compararlo con el trabajado en el presente documento. En primer término, debe ser tenido en cuenta que ambos índices presentan una diferencia conceptual importante, el IPM de Alkire permite realizar medidas “absolutas” de pobreza, es decir, considera una definición de pobreza mediante sus dimensiones que se puede aplicar a cualquier hogar del cual se desea conocer su estatus de pobreza. Por otro lado el IPM planteado en el presente trabajo es una medida “relativa” de pobreza, dado que la metodología del AFM hace que el índice no esté definido únicamente por las variables y las dimensiones que constituyen sino por los hogares considerados a la hora de realizar el análisis. Esto puede resultar una limitante seria si se pretende usar el índice para la comparación entre distintos grupos de hogares, por ejemplo, entre distintos países.

Teniendo en cuenta esta diferencia, un aspecto negativo es la menor interpretabilidad directa del índice respecto al IPM de Alkire (que se puede interpretar como una proporción de privaciones) y como resultado de esto la mayor complejidad que presenta para fijar un punto de corte del mismo a partir del cual un hogar es considerado multidimensionalmente pobre.

Por otro lado, el IPM trabajado tiene como aspecto a favor la posibilidad de incluir un número elevado de variables, que mediante la metodología del AFM se tratan de resumir con la menor pérdida de información posible. Adicionalmente, el AFM permite ver las relaciones entre las categorías de las variables y entre el índice final y los resultados intermedios que se obtienen producto de la primera fase del análisis.

A futuro deberán ser puestas a discusión las dimensiones, en particular con especialistas del estudio de la pobreza, para contar con un conjunto que sea más representativo del fenómeno. También se deberá profundizar en el estudio de la diferencia entre las regiones

tratadas y las diferencias entre los dos acercamientos al AFM estudiadas. Otro aspecto a tomar en cuenta del IPM obtenido es que en la aplicación realizada no se puede ponderar arbitrariamente las dimensiones, lo cual puede resultar de interés. Como comentario general a toda medida de resumen, debe tenerse en cuenta que la simplificación realizada implicará una pérdida de información. Deberá considerarse si resulta atractivo trabajar en una medida de resumen cuyo resultado final sea un único valor, por ejemplo, un conjunto de medidas a partir de los resultados del AFM.

Referencias

- [1] O. Aires and F. Di Landri, “Estimación de la pobreza por el método del ingreso, año 2017,” Instituto Nacional de Estadística, Montevideo, Uruguay, Tech. Rep., Abril 2018.
- [2] S. Alkire and J. Foster, “Counting and multidimensional poverty measurement,” Oxford Poverty & Human Development Initiative (OPHI), Oxford, Inglaterra, Tech. Rep. 32, Diciembre 2009.
- [3] B. Escofier and J. Pagès, “Multiple factor analysis (afmult package),” *Computational Statistics & Data Analysis*, vol. 18, pp. 121–140, 1994.
- [4] S. Alkire, U. Kanagaratnam, and N. Suppa, “The global multidimensional poverty index (mpi): 2018 revision,” Oxford Poverty and Human Development Initiative (OPHI), Oxford, Inglaterra, Tech. Rep. 47, Setiembre 2018.
- [5] C. Collicelli and M. Valerii, “A new methodology for comparative analysis of poverty in the mediterranean: A model for differential analysis of poverty at a regional level via the principal components analysis,” Fondazione CENSIS, Roma, Italia, Tech. Rep., Agosto 2000.
- [6] A. Camardiel, L. Vásquez, and G. Ramírez, “Una propuesta para la construcción de un índice sintético de pobreza,” *Revista Venezolana de Análisis de Coyuntura*, vol. VI, no. 1, pp. 121–142, 2000.

-
- [7] J. J. Calvo, V. Borrás, W. Cabella, P. Carrasco, H. De los Campos, M. Koolhaas, M. Macadar, M. Nathan, N. nez, I. Pardo, M. Tenenbaum, and V. C., “Atlas socio-demográfico y de la desigualdad del uruguay: Las necesidades básicas insatisfechas a partir de los censos 2011,” Instituto Nacional de Estadística, Montevideo, Uruguay, Tech. Rep., Julio 2013.
- [8] M. Colafranceschi, M. Peyrou, and M. Sanguinetti, “Pobreza multidimensional en uruguay: una aplicación de técnicas multivariadas,” Master’s thesis, Facultad de Ciencias Económicas y Administración, UdelaR, Montevideo, Uruguay, Febrero 2009.
- [9] L. Nalbarte, F. Massa, and S. Altmark, “Pobreza multidimensional utilizando análisis factorial,” 2014.
- [10] J. Pagès, *Multiple Factor Analysis by Example Using R*. Boca Raton, Florida: CRC Press, 2015.
- [11] O. Aires, F. Di Landri, and C. Medero, “Principales resultados de la encuesta nacional de gastos e ingresos de los hogares (2016-2017),” Instituto Nacional de Estadística, Montevideo, Uruguay, Tech. Rep., Febrero 2020.
- [12] S. Alkire, “Choosing dimensions: the capability approach and multidimensional poverty,” Oxford Poverty and Human Development Initiative (OPHI), Oxford, Inglaterra, Tech. Rep. 88, Agosto 2007.
- [13] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2021. [Online]. Available: <https://www.R-project.org/>
- [14] H. Wickham, M. Averick, J. Bryan, W. Chang, L. D. McGowan, R. François, G. Grolemund, A. Hayes, L. Henry, J. Hester, M. Kuhn, T. L. Pedersen, E. Miller, S. M. Bache, K. Müller, J. Ooms, D. Robinson, D. P. Seidel, V. Spinu, K. Takahashi, D. Vaughan, C. Wilke, K. Woo, and H. Yutani, “Welcome to the tidyverse,” *Journal of Open Source Software*, vol. 4, no. 43, p. 1686, 2019.
- [15] S. Lê, J. Josse, and F. Husson, “FactoMineR: A package for multivariate analysis,” *Journal of Statistical Software*, vol. 25, no. 1, pp. 1–18, 2008.

- [16] J. Blanco, *Introducción al Análisis Multivariado: Teoría y aplicaciones a la realidad latinoamericana*. Montevideo, Uruguay: IESTA, FCEA, UdelaR, 2006.



FACULTAD DE
CIENCIAS ECONÓMICAS
Y DE ADMINISTRACIÓN

IESTA 80 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 4/22