



# Resolución Ministerial

N° 107-2015-MIDIS

Lima, 08 MAYO 2015

## VISTOS:

El Informe N° 88-2015-MIDIS/VMPES/DGGU de la Dirección General de Gestión de Usuarios; y el Informe N° 017-2015-MIDIS-VMPES/DGGU-DO de la Dirección de Operaciones; y.

## CONSIDERANDO:

Que, mediante Ley N° 29792, se creó el Ministerio de Desarrollo e Inclusión Social - MIDIS, determinándose su ámbito, competencias, funciones y estructura orgánica básica; asimismo, se estableció que el sector Desarrollo e Inclusión Social comprende a todas las entidades del Estado, de los tres niveles de gobierno vinculados con el cumplimiento de las políticas nacionales en materia de promoción del desarrollo social, la inclusión y la equidad;

Que, el artículo 9 de la Ley N° 29792 dispone que el MIDIS desarrolla como función exclusiva el diseño y gestión de los registros y base de datos de los programas sociales, de identificación, selección y registro de beneficiarios y otros registros que se creen;

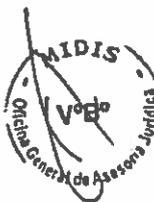
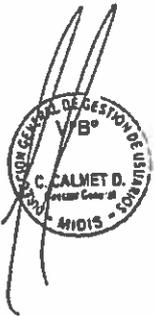
Que, mediante Resolución Ministerial N° 399-2004-PCM se creó el Sistema de Focalización de Hogares - SISFOH, con el propósito de proveer información a los programas sociales para la identificación y selección de sus beneficiarios;

Que, mediante Resolución Ministerial N° 320-2010-PCM, se aprobó la Metodología de Cálculo del Índice de Focalización de Hogares, la cual viene siendo utilizada por la Unidad Central de Focalización (UCF);

Que, el artículo 26 de la Ley N° 29951, Ley de Presupuesto del Sector Público para el Año Fiscal 2013, dispuso que la UCF, a cargo de la operación del SISFOH, se encuentra bajo el ámbito de la Dirección General de Gestión de Usuarios del Ministerio de Desarrollo e Inclusión Social, con carácter permanente, correspondiéndole a dicha unidad, generar y administrar la información del Padrón General de Hogares (PGH) bajo estándares de calidad, seguridad y confidencialidad, así como certificar la Clasificación Socioeconómica de los potenciales usuarios ante los agentes responsables de la administración de los programas sociales y de subsidios del Estado que se ejecuten bajo criterios de focalización individual;

Que, según el citado artículo, compete al Ministerio de Desarrollo e Inclusión Social dictar los lineamientos, criterios, reglas y procedimientos para la focalización individual que contribuya a mejorar la equidad y eficiencia en la asignación de los recursos públicos en beneficio de aquellos grupos poblacionales priorizados;

Que, conforme con el artículo 50 del Reglamento de Organización y Funciones del Ministerio de Desarrollo e Inclusión Social, aprobado por Decreto Supremo N° 011-2012-MIDIS, la Dirección de Operaciones es la unidad orgánica de la Dirección General de Gestión de



Usuarios encargada de diseñar y desarrollar la plataforma operativa para el funcionamiento del sistema de Información de usuarios, programas y/o proyectos del Ministerio, que permite contar con información prioritaria para el Sector;

Que, en el marco de dicha competencia, mediante Resolución Ministerial N° 277-2014-MIDIS, se autorizó la actualización del algoritmo de cálculo del Índice de Focalización de Hogares, sobre la base de la metodología vigente, con las variables de la Encuesta Nacional de Hogares (ENAH) 2010 - 2012;

Que, recientemente el Instituto Nacional de Estadística e Informática (INEI) ha publicado la Información correspondiente a la Encuesta Nacional de Hogares (ENAH) 2014;

Que, en dicho contexto, mediante los documentos de Vistos, la Dirección de Operaciones y la Dirección General de Gestión de Usuarios proponen reemplazar la Metodología de Cálculo del Índice de Focalización de Hogares, por una nueva, denominada "Metodología para la Determinación de la Clasificación Socioeconómica";

Que, las citadas Direcciones, indican que la nueva metodología debería aplicarse únicamente a las nuevas solicitudes de Clasificación Socioeconómica, de actualización de Clasificación Socioeconómica y a las de reevaluación, así como también a todas aquellas solicitudes que se encuentran en trámite, tomando en cuenta que la información con la que se determinó la Clasificación Socioeconómica de los hogares actualmente registrados en el PGH, corresponde a un contexto económico y social diferente al de las estimaciones realizadas en la presente propuesta;

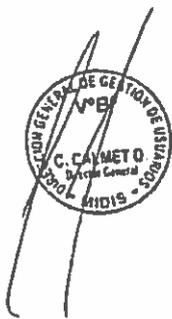
Que, asimismo, el artículo 26 de la Ley N° 29951, Ley de Presupuesto del Sector Público para el Año Fiscal 2013, indica que la incorporación de usuarios a los programas sociales y de subsidios del Estado deberá tomar en cuenta la Clasificación Socioeconómica de las personas que contiene el PGH; en razón de ello, se sustituyen los términos de elegible y no elegible por el resultado de la Clasificación Socioeconómica;

Que, atendiendo a lo expuesto y considerando las competencias y funciones del MIDIS, establecidas en las normas antes reseñadas, se considera pertinente aprobar la Metodología para la Determinación de la Clasificación Socioeconómica, la cual reemplaza a la Metodología de Cálculo del Índice de Focalización de Hogares;

De conformidad con lo dispuesto en la Ley N° 27972, Ley de Creación, Organización y Funciones del Ministerio de Desarrollo e Inclusión Social, y su Reglamento de Organización y Funciones, aprobado por Decreto Supremo N° 011-2012-MIDIS;

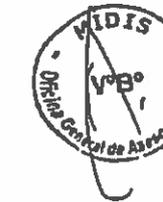
**SE RESUELVE:**

**Artículo 1.-** Aprobar la "Metodología para la Determinación de la Clasificación Socioeconómica", cuyo texto, en anexo, forma parte integrante de la presente resolución.





# Resolución Ministerial



**Artículo 2.-** Dejar sin efecto la Metodología de Cálculo del Índice de Focalización de Hogares, aprobada mediante Resolución Ministerial N° 320-2010-PCM.

**Artículo 3.-** Disponer que la Metodología para la Determinación de la Clasificación Socioeconómica sea de aplicación únicamente para las nuevas solicitudes de Clasificación Socioeconómica, de actualización de Clasificación Socioeconómica y de reevaluación, así como a aquellas solicitudes que se encuentran en trámite.

**Artículo 4.-** Disponer que la Dirección General de Gestión de Usuarios, cada vez que el Instituto Nacional de Estadística e Informática publique información sobre una nueva Encuesta Nacional de Hogares, evalúe la pertinencia de presentar una propuesta de actualización de las variables de la Metodología para la Determinación de la Clasificación Socioeconómica.

**Artículo 5.-** Disponer la publicación de la presente resolución y su anexo en el portal Institucional del Ministerio de Desarrollo e Inclusión Social ([www.mdis.gob.pe](http://www.mdis.gob.pe)), en la misma fecha de publicación de la presente resolución.

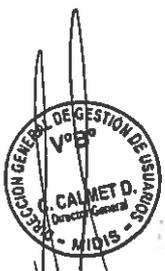
**Regístrese, comuníquese y publíquese.**

.....  
Paola Bustamante Suárez  
MINISTRA DE DESARROLLO E INCLUSIÓN SOCIAL



**Ministerio de Desarrollo e Inclusión Social**  
**Dirección General de Gestión de Usuarios**

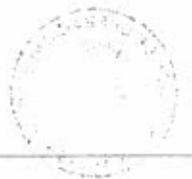
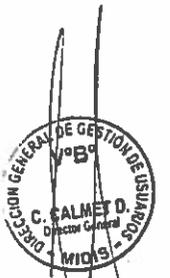
**METODOLOGÍA PARA LA DETERMINACIÓN DE LA  
CLASIFICACIÓN SOCIOECONÓMICA**



Mayo 2015

## Contenido

I.	Introducción .....	3
II.	Procedimiento de Determinación de Elegibilidad Vigente .....	4
III.	Metodología para la Determinación de la Clasificación Socioeconómica .....	6
A.	Consideraciones Generales .....	6
B.	Consideraciones Específicas .....	7
C.	Propuesta de Metodología para la Determinación de la Clasificación Socioeconómica .....	9
IV.	Metodología de Cálculo del IFH .....	10
A.	Selección de Variables mediante Análisis de Correlación y Dinámica Temporal .....	10
B.	Análisis de Componentes Principales - PCA .....	13
C.	Determinación de Puntos de Cortes para la CSE .....	14
D.	Criterios de Evaluación de la Capacidad Predictiva .....	15
V.	Índice de Focalización de Hogares (Algoritmo) y Capacidad Predictiva .....	17
A.	Selección de variables mediante análisis de correlación y dinámica temporal .....	17
B.	Análisis de Componentes Principales - PCA .....	18
C.	Determinación de Puntos de Cortes para la CSE .....	19
D.	Evaluación de la Capacidad Predictiva .....	20
VI.	Conclusiones .....	21
VII.	Anexos .....	23



## I. Introducción

La identificación y selección de beneficiarios es un reto constante para los programas sociales debido a que no existe una manera automática y única de establecer cuándo un hogar es calificado como elegible para acceder a un programa social. Esta tarea resulta particularmente compleja si consideramos el contexto de alta informalidad de la economía peruana, en la que la recolección de datos sobre la riqueza o ingresos de los hogares resulta tanto costosa como propensa a los problemas de errores de medición (i.e. sub-reporte de los sectores de altos ingresos).

Es así que el 2004 se introduce el Sistema de Focalización de Hogares (SISFOH) con el propósito de proveer información a los programas sociales para la identificación y selección de los potenciales usuarios de los programas sociales. Para ello, el SISFOH sigue a partir del 2010 una metodología que permite determinar de manera eficiente la Clasificación Socioeconómica (CSE) de los hogares, en un contexto como el mencionado, en el que el nivel de los ingresos es difícil de captar por medios verificables y, por tanto, debe ser aproximado a través de metodologías apropiadas para dicho fin.

La metodología vigente para la determinación de la elegibilidad de los hogares - en adelante Metodología MEF-2010 - fue diseñada utilizando fuentes de información del año 2009 y, por tanto, representa las condiciones de vida existentes en el país en dicho periodo de tiempo. Ciertamente, en el transcurso de los años, la sociedad peruana ha experimentado profundos cambios demográficos, sociales y económicos, que han impactado en el mercado de trabajo y las condiciones de vida de la población.

En detalle, la calidad de vida de las personas, en atención a los resultados de la Encuesta Nacional de Hogares (ENAHO), presenta mejoras importantes respecto al 2009, principalmente en el segmento de la población más pobre y del ámbito rural. El gasto e ingreso real per-cápita promedio en el ámbito rural se incrementa en 8 y 22 puntos porcentuales (frente a 12 y 9 puntos porcentuales de incremento en el ámbito urbano- incluyendo Lima Metropolitana) y la cobertura de servicios básicos de agua de red pública y electricidad para el alumbrado de la vivienda para los hogares en condición de pobreza se extiende en 17 puntos porcentuales. Una cifra similar se obtiene para el incremento de la cobertura en el abastecimiento de agua en los hogares en condiciones de pobreza extrema y una equivalente a 13 puntos porcentuales en el acceso a la electricidad para el mismo grupo.

Esto genera la necesidad de actualizar los parámetros empleados para el cálculo de los puntos de corte y evaluar la pertinencia de las herramientas disponibles. Atendiendo a las circunstancias mencionadas, la Dirección General de Gestión Usuarios, en el marco de su competencia, identifica la necesidad de establecer el procedimiento de determinación de Clasificación Socioeconómica sobre la base de la ENAHO 2014, que reemplaza la actual Metodología de Cálculo del Índice de Focalización de Hogares.

De esta forma, el documento describe la nueva metodología de determinación de la Clasificación Socioeconómica. Para tal efecto se analizan las interrelaciones entre diversas variables relacionadas con el hogar y la vivienda, así como con la educación y ocupación, entre otras, utilizando la información de la ENAHO que comprende el periodo entre los años 2007 a 2014, de forma tal que se determinen variables comunes para el planteamiento y validación del modelo resultante de la metodología. Asimismo, cabe señalar que se considera como elemento fundamental la ubicación geográfica del hogar y específicamente la ruralidad del ámbito específico de residencia del hogar.

El resto del documento se ordena de la siguiente forma, el capítulo II presenta la propuesta metodológica, el capítulo III detalla el método de estimación del Índice de Focalización de Hogares (IFH), el capítulo IV presenta los indicadores de resultado de la calidad de predicción del IFH propuesto y del algoritmo (el que incluye además el uso de remuneraciones y consumo de agua y electricidad). El capítulo V cierra el documento presentando las conclusiones y recomendaciones que resultan de la presente metodología.



## II. Procedimiento de Determinación de Elegibilidad Vigente

La Unidad Central de Focalización (UCF), como administrador del PGH, determina la Clasificación Socioeconómica (CSE) de los hogares sobre la base de la Metodología MEF-2010. En conformidad con la Resolución Ministerial N° 320-2010-PCM, que aprobó la Metodología de Cálculo del Índice de Focalización de Hogares y la Ficha Socioeconómica Única (FSU), el procedimiento que emplea la UCF para determinar si una persona es elegible para un programa social específico sigue tres pasos:

- (i) Primer paso: A partir de la información de los registros administrativos que contienen las planillas del sector público y privado, se identifica a los perceptores que superan el punto de corte de S/. 1,500 Nuevos Soles, calificándolos como “no pobre” y, por tanto, no elegibles para los programas que así lo consideren. Este punto de referencia se fija teniendo en consideración un número de miembros del hogar de 5 o menos y la línea de pobreza monetaria del 2009 de S/ 300 Nuevos Soles, cifra oficial calculada por el Instituto Nacional de Estadística e Informática- INEI. En este sentido, este primer paso filtra a los hogares con ingresos por planilla por encima de los S/. 1,500 Nuevos Soles.
- (ii) Segundo paso: A partir de un Índice de Focalización de Hogares (IFH) se evalúa la condición socioeconómica del hogar sobre la base de indicadores de calidad de vida que se recogen en la Ficha Socioeconómica Única (FSU); y umbrales que corresponden con determinado costo admisible de error de focalización y que varían de acuerdo al departamento, el ámbito urbano/ rural y la región natural en la que reside el hogar. El modelo es estimado en base a la información de la Encuesta Nacional de Hogares (ENAH) del 2009.
- (iii) Tercer paso: Se basa en el consumo de agua o electricidad tomando como punto de corte S/. 20 y S/. 25 Nuevos Soles respectivamente. Cifras calculadas en base a la mediana del consumo de la población en situación de pobreza que tiene acceso a dicho servicio de la ENAHO 2009. A la salida de este proceso se determinaba la elegibilidad o la no elegibilidad a programas sociales.

En correspondencia con los criterios establecidos en la Metodología MEF-2010, el siguiente Gráfico 1 resume el procedimiento de determinación de elegibilidad:

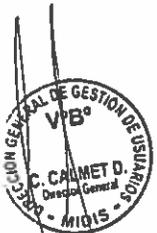
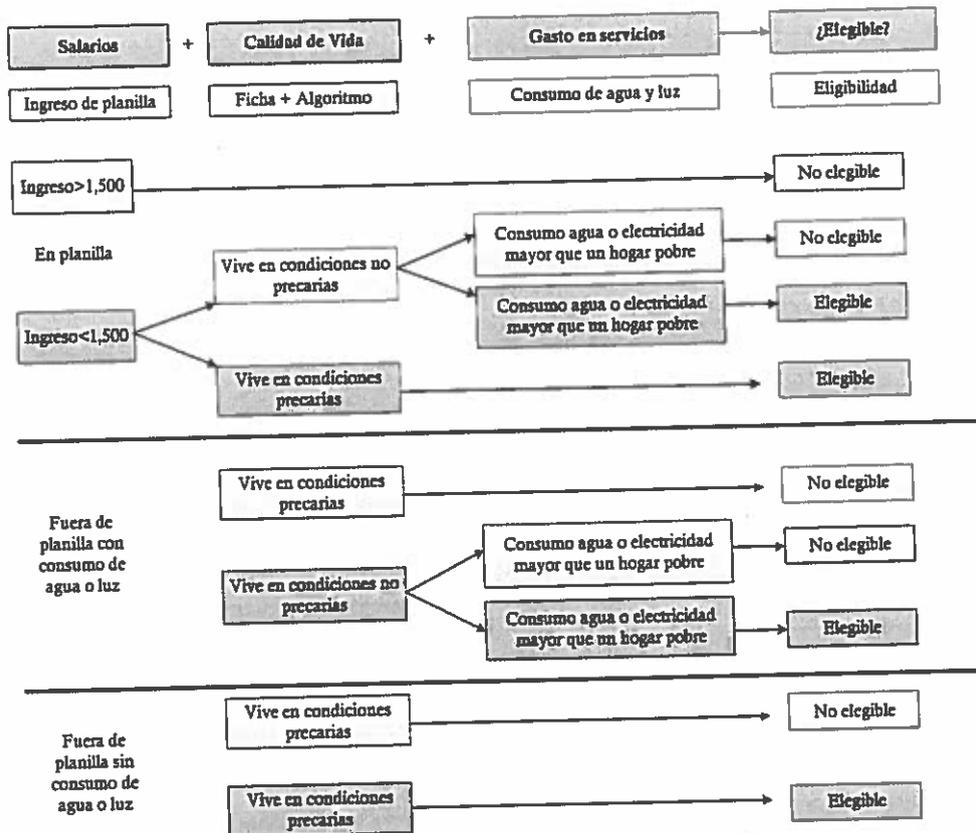


Gráfico 1 Metodología para determinar la elegibilidad



Fuente: Resolución Ministerial N° 320-2010-PCM

Nota: En el documento original se observa la frase "consumo agua o electricidad mayor que un hogar pobre" en los cuadros señalados. Al tratarse de un error material, se implementó como "consumo agua o electricidad menor que un hogar pobre".

Como se mencionó previamente, la Metodología MEF- 2010 fue diseñada utilizando fuentes de información del año 2009 y representa, por tanto, las condiciones de vida existentes en el país en ese periodo de tiempo. Sin embargo, en el transcurso de los años la sociedad peruana ha experimentado cambios demográficos, sociales y económicos, que han impactado en el mercado de trabajo y las condiciones de vida de la población. En consecuencia, es imperativo que la Metodología MEF-2010 actualice los parámetros empleados para el cálculo de los puntos de corte; y de esta manera determinar la pertinencia de las herramientas disponibles.

Al respecto cabe mencionar que mediante Resolución Ministerial N° 277-2014-MIDIS, se autorizó la actualización del algoritmo de cálculo del IFH, sobre la base de la metodología vigente, con las variables de la encuesta ENAHO 2010 - 2012.



### III. Metodología para la Determinación de la Clasificación Socioeconómica

#### A. Consideraciones Generales

A partir del análisis efectuado a la metodología vigente, se consideran necesarias la incorporación de las siguientes mejoras:

- *Medición que enfatiza la atención en el hogar y su relación con el territorio*

Un sistema de focalización debe considerar en su diseño y análisis el rol del hogar, en tanto espacio en el que se expresan las diversas etapas del ciclo de vida y como recepto natural de las intervenciones del Estado. Por tanto, se realiza un giro en el énfasis a la hora de medir los recursos económicos. Pasando de considerar al individuo como generador de ingresos desvinculado de una unidad familiar a un enfoque de hogar. De esta forma se consideran los ingresos de todos los perceptores miembros del grupo familiar. Asimismo, se incorpora la heterogeneidad entre hogares en función de las características de los miembros del hogar. Es así, que se considera oportuno priorizar la información de recursos del hogar en su versión per cápita en lugar de utilizar montos totales. Esto con la finalidad de atender las necesidades diferenciadas de los hogares en función del tamaño del grupo familiar y reducir así las potenciales distorsiones generadas.

Asimismo, a fin de incorporar las diferencias en las dinámicas territoriales, la evaluación de los pasos i y iii de la Metodología MEF-2010 se desagrega en tres grupos: Lima Metropolitana, Resto Urbano y Rural. La diferenciación por el criterio de ruralidad, considera que la baja densidad de la población y alta dispersión de las viviendas del ámbito rural repercute y condiciona la capacidad de los hogares para aprovechar las oportunidades económicas. Asimismo, se incorpora un gradiente de urbanidad adicional considerando a Lima Metropolitana como un separado del resto urbano, en virtud del mayor dinamismo de la actividad económica, mejor provisión de servicios públicos y mercados más desarrollados.

Por su parte, para fines del paso ii, respecto al cálculo del IFH, al ser necesaria una mayor precisión en los cálculos, se desagrega el territorio nacional en quince conglomerados<sup>1</sup> que articulan información geográfica y social, sobre la base de patrones de consumo diferenciados. En la siguiente sección se presentara a detalle dicha distinción.

- *Ampliación de herramientas para el análisis del desempeño de los instrumentos de focalización*

Se amplían las metodologías para la selección de umbrales o puntos de corte: Se evalúan umbrales que optimizan diferentes funciones de pérdida. Las metodologías de cálculo de umbrales propuestas son: i) La minimización del error de focalización, que escoge un umbral del indicador que minimiza el error de focalización-definido como la suma ponderada de los errores de exclusión y filtración; y, ii) el Umbral de Exclusión del 2%, que selecciona el umbral del indicador que excluye al 2% de las personas en situación de pobreza. Una explicación detallada se encuentra en el Capítulo V.

Asimismo, cabe precisar que si bien la validez intrínseca de los instrumentos empleados depende de la capacidad de obtener un resultado concreto (pobre o no pobre) en función de la real condición de pobreza del hogar (i.e. la sensibilidad y la especificidad); el evaluador carece de esta

<sup>1</sup> Costa Norte Rural, Costa Norte Urbana, Costa Centro Rural, Costa Centro Urbana, Costa Sur Rural, Costa Sur Urbana, Sierra Norte Rural, Sierra Norte Urbana, Sierra Centro Rural, Sierra Centro Urbana, Sierra Sur Rural, Sierra Sur Urbana, Selva Rural, Selva Urbana y Lima Metropolitana.

## Metodología para la Determinación de la Clasificación Socioeconómica

información *a priori* y, por tanto, pierde relevancia en la práctica.<sup>2</sup> Por ello, la incorporación de información sobre los valores predictivos asociados a los puntos de corte de cada instrumento prioriza información sobre la precisión del resultado concreto de la metodología (pobre o no pobre), el que depende de la tasa de pobreza observada. Al respecto, se calcula para cada umbral lo proveniente de las metodologías previamente mencionadas, la probabilidad de que el hogar sea realmente un hogar pobre (o no) ante un resultado concreto de la metodología (valores predictivos) es más certera.

En esta lógica, se extiende el cálculo de los valores predictivos a la secuencia de pasos dentro de la metodología, las que se evaluarán en función al aporte en la predicción de la pobreza. Puesto que la precisión del proceso en conjunto depende no solo de los instrumentos usados, sino de cómo y en qué contexto son usados, en ese sentido se parte de la actualización del algoritmo de cálculo del IFH, para evaluar la pertinencia de la implementación de los pasos asociados a la incorporación de las bases de consumo de servicios básicos agua y luz y de los registros administrativos de las planillas públicas y privadas.

En este sentido, se analizan diferentes secuencias de procedimientos considerando el ámbito de aplicación y si se procede a la evaluación de forma simultánea (en paralelo) o en función de los resultados de las fases previas (en serie). El arreglo de instrumentos que configuren finalmente la metodología se determina en función del aporte a la predicción, medido como el cambio marginal de los valores predictivos. Información sobre valores predictivos de los arreglos finales deberá ser incluida al momento de procesar la información con la que la UCF determinará la Clasificación Socioeconómica.

### - Mejora en la precisión de la terminología

Se propone reemplazar la denominación del instrumento como Metodología para la Determinación de la Clasificación Socioeconómica, que sustituye a la Metodología MEF-2010. Esto responde a la Directiva N° 1-2015-MIDIS<sup>3</sup>, que regula en el numeral 6.1 la Determinación de la Clasificación Socioeconómica en concordancia con las funciones que debe cumplir la UCF en atención a su normativa vigente. Adicionalmente, se incorpora definiciones operacionales que permiten que el procedimiento de medición sea lo suficientemente explicativo como para ser realizado por diferentes personas de forma independiente y se presenta un glosario actualizado que incorpora los cambios en los procesos desde la metodología anterior.

### B. Consideraciones Específicas

A fin de mejorar la precisión del proceso en conjunto, y teniendo en cuenta para ello las recomendaciones metodológicas de los párrafos precedentes, se propone la siguiente Metodología para la determinación de la Clasificación Socioeconómica:

#### - Paso 1: Ingresos Laborales proveniente de Planillas Públicas y Privados

En conformidad con los criterios expuestos en las consideraciones generales, se emplea como indicador el ingreso familiar per-cápita, definido como la suma de los ingresos que se derivan de

<sup>2</sup> El teorema de Bayes y la inferencia bayesiana presentan un marco estadístico para recalcularse la probabilidad respecto de un elemento de un conjunto que contiene una determinada característica en base a: (i) la probabilidad inicial de poseer dicha característica y (ii) los errores conocidos de los instrumentos usados para determinar la característica. La probabilidad final se denomina valor predictivo y el cálculo de los errores de los instrumentos aplicados a este marco estadístico se denomina cálculo de verosimilitud. En ese sentido, es posible ver a la pobreza como un atributo de un individuo/hogar u otra unidad en un conjunto más grande, cuya probabilidad inicial corresponde a la tasa de pobreza de su conjunto, y en la cual esta probabilidad de ser pobre es modificada por las mediciones de umbrales e índices de focalización. Un elemento con una probabilidad recalculada cercana a 100% entonces es muy probable de ser pobre, mientras que el 0% indica con mucha certeza que es no pobre.

<sup>3</sup> Aprobada por Resolución Ministerial N° 23-2015-MIDIS, publicada en el diario oficial "El Peruano", el 22 de enero del 2015



### Metodología para la Determinación de la Clasificación Socioeconómica

las competencias laborales del total de miembros del hogar en edad de trabajar ocupados dividido entre el número de miembros del hogar. La cifra obtenida se enfrenta a los cortes actualizados para las planillas públicas y privadas, para Lima Metropolitana, Resto Urbano y Rural, en función al promedio de las líneas de pobreza para cada ámbito calculadas por el INEI a partir de la ENAHO 2014, en base a la metodología de medición de pobreza actualizada.<sup>4</sup>

Para la selección de umbrales se tomó en cuenta que estos instrumentos son usados como primera línea de detección de pobreza del hogar solicitante, por lo cual se requiere altos valores de sensibilidad que mejoren el rendimiento del Índice de focalización de hogares para la clasificación en etapas posteriores. Por ello no solo es necesario evaluar los valores de error de inclusión / exclusión (y sus contrapartes sensibilidad / especificidad) sino que se evaluó el rendimiento predictivo para las frecuencias de pobreza que se esperaba se encuentre en campo, tratando que la predicción negativa no baje del 90%.

Por su parte, la información de Planillas Públicas y Privadas presenta los ingresos laborales de trabajadores del sector formal de la economía, la que según cifras del 2012 incorpora sólo al 26% de la Población Económicamente Activa (PEA): 40%, 33% y 5% de la PEA de Lima Metropolitana, ámbito urbano y rural, respectivamente.<sup>5</sup> Esto añade una dificultad a la hora de calcular los umbrales ya que al considerar únicamente a los formales, se contrasta grupos de hogares pobres y no pobres menos disímiles que los que resultarían de un contraste ejecutado en la población total. Es por ello que se podría castigar a los pobres en el sector formal, ya que es mucho más difícil discernir entre ambos grupos. Para evitar esta distorsión en los ámbitos rural y resto urbano se emplea un valor de corte mayor a la línea de pobreza, puesto que esta presenta una sensibilidad baja.

Asimismo, en atención a la alta rotación entre la formalidad y la informalidad en el mercado laboral peruano, se debe considerar que estos ingresos pueden reflejar contingencias de los hogares y no necesariamente condiciones permanentes. Es por ello que se considera un mínimo de datos de tres meses en el periodo de semestral previo para que esta información se incorpore en el proceso.

Teniendo en consideración lo expuesto, el umbral para Lima Metropolitana se fija en S/.383 Nuevos Soles, para el resto urbano en S/.337 Nuevos Soles y para el ámbito rural en S/.320 Nuevos Soles.

#### - Paso 2: Información sobre el Consumo en servicios de Agua y Electricidad

La información de consumo en servicios de agua y electricidad presenta los consumos en dichos servicios para los hogares que cuentan con acceso y reportaron algún consumo. El subgrupo de hogares que logra abastecerse de agua dentro de la vivienda a través de la red pública y cuenta con alumbrado eléctrico en la vivienda representa para el 2013 el 76% y 90% del total de hogares. Al desagregar la información por ámbito, se obtiene que el 96%, 97% y 70% de los hogares cuenta con alumbrado en la vivienda en Lima Metropolitana, en el resto urbano y en el ámbito rural, respectivamente. Mientras que para el caso del abastecimiento de agua, las cifras son 87%, 81% y 53% en Lima Metropolitana, resto urbano y ámbito rural, respectivamente.

Para el presente año no se ha usado los valores de consumo de agua debido a fluctuaciones no predecibles en los montos de consumo y la poca disponibilidad de estos valores con respecto a los consumos de electricidad. En conformidad con los criterios expuestos en las consideraciones generales, se emplea como indicador el consumo per-cápita del hogar en los servicios de

<sup>4</sup> Los ejercicios con la versión per-cápita de los ingresos en relación a los montos agregados presenta una mejor capacidad de predicción, sustentada en áreas debajo de la curva ROC de mayor valor. La curva ROC es una representación gráfica de los errores de predicción (sensibilidad, especificada) que indica el éxito a la hora de predecir la condición (i.e. la pobreza). Esta varía en el rango de 0% a 100%, en donde 0% indica una predicción completamente imperfecta y 100% una predicción completamente perfecta.

<sup>5</sup> INEI (2013). "Producción y Empleo Informal en el Perú. Cuenta Satélite de la Economía Informal 2007-2012". Lima: INEI.

## Metodología para la Determinación de la Clasificación Socioeconómica

abastecimiento de agua dentro de la vivienda y el alumbrado eléctrico de red pública; definidos como el consumo en dichos servicios divididos entre el número de miembros del hogar.

En cuanto a la selección de umbrales, se evalúa el punto de corte vigente - calculado según el consumo promedio entre los hogares pobres- y adicionalmente los umbrales estimados a partir de las metodologías señaladas. Para ello, se cuantifica el aporte marginal de cada umbral en el valor predictivo negativo en el ámbito rural (en donde es menos creíble un resultado negativo o no pobre) y en el valor predictivo positivo en el ámbito urbano (en donde es menos creíble un resultado positivo o pobre), respecto a los valores predictivos positivos y negativos obtenidos de emplear únicamente el IFH.

Bajo este criterio se optó por los valores de consumo de electricidad per cápita de S/. 18, S/. 10.1 y S/. 9.2 para los ámbitos de Lima Metropolitana, urbano y rural respectivamente.

### - Paso 3: Cálculo del Índice de Focalización de Hogares

Es preciso señalar, que la DGGU, en el marco de sus funciones, dio inicio al proceso de actualización de la metodología con la actualización del IFH, el mismo que fue aprobado mediante Resolución Ministerial N° 277-2014-MIDIS, del 02 de diciembre del año 2014. Para este efecto, se actualiza el algoritmo de cálculo del IFH, sobre la base de la metodología vigente, con las variables de la encuesta ENAHO 2010-2012.

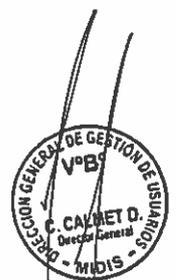
El Informe N° 027-2014-MIDIS-VMPES/DGGU/DO brinda evidencia de la baja exactitud para la detección de la pobreza del IFH anterior. Estos fenómenos suelen ser comunes con las metodologías usadas, por lo cual se requiere de su actualización periódica. En concordancia con ello, se presenta en el Capítulo IV del presente documento la propuesta metodológica para el Cálculo del Índice de Focalización a la luz de la publicación de la ENAHO 2014 (en adelante IFH-2015). Para ello se incorporan las consideraciones generales expuesta anteriormente.

### C. Propuesta de Metodología para la Determinación de la Clasificación Socioeconómica

En consecuencia con el análisis precedente realizado sobre la base de las consideraciones metodológicas expuestas, se propone la Metodología para la Determinación de la Clasificación Socioeconómica<sup>6</sup>:

- (i) Evaluación del Nivel de Ingresos: Se contrasta el ingreso neto por planilla pública o privada per cápita, incluyendo información de todos los perceptores del hogar, con el punto de corte de S/. 383 Nuevos Soles para Lima Metropolitana, S/. 337 Nuevos Soles para el resto urbano y S/. 320 Nuevos Soles para el ámbito rural. Aquellos hogares que sobrepasen estos montos serán considerados como no pobres. Los hogares cuyo ingreso sea igual o menor que el umbral, o en los que no se cuente con información seguirán al siguiente paso.
- (ii) Evaluación del gasto en consumo de electricidad: Se compara el consumo de electricidad considerando los puntos de corte de S/. 18, S/. 10.1 y S/. 9.2 para los ámbitos de Lima Metropolitana, urbano y rural respectivamente. Los hogares con un consumo per cápita superior son considerados como no pobres. Aquellos con consumo menor o igual, o en los que no se cuente con información seguirán al siguiente paso.
- (iii) Evaluación de la calidad de vida del hogar: El IFH-2015 será aplicado a los hogares que se encuentren por debajo de los umbrales anteriores o que no hayan tenido información previa. Este resultado, interpretado como No Pobre, Pobre o

<sup>6</sup> En el Anexo B, se presenta el Glosario de términos.

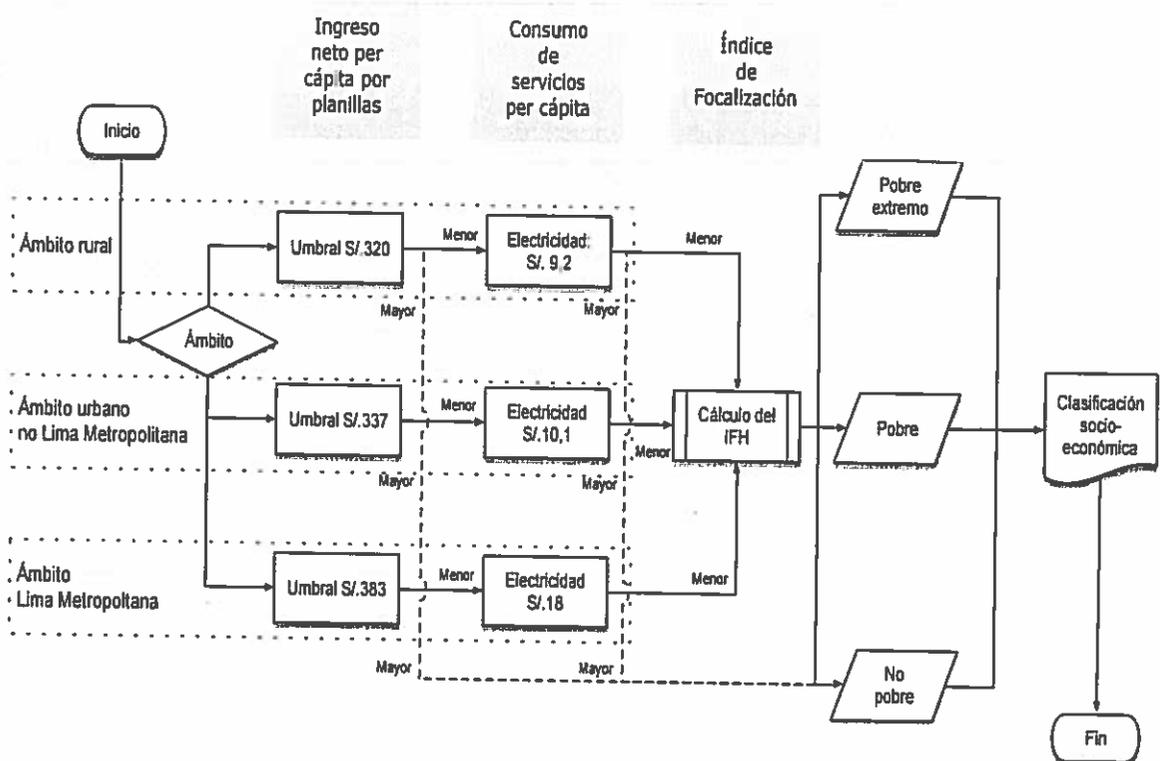


### Metodología para la Determinación de la Clasificación Socioeconómica

Pobre Extremo según los puntos de cortes determinados por la Dirección General de Gestión de Usuarios, será publicado en el Padrón General de Hogares (PGH), así como la asignación del estrato que le corresponda.

Así, el Gráfico 2 resume la propuesta:

**Gráfico 2. Metodología para la Determinación de la Clasificación Socioeconómica**



Elaboración: DO-DGGU

#### IV. Metodología de Cálculo del IFH

La metodología que se emplea para la construcción del Índice de Focalización de Hogares (IFH-2015) se conforma por la siguiente secuencia de procedimientos: (i) Construcción de la base de datos, en función a la información proveniente de la ENAHO 2007-2014, la que anualmente es recogida por el Instituto Nacional de Estadística e Informática-INEI; (ii) selección de variables, empleando como criterio la correlación que éstas exhiben en relación a medidas de bienestar del hogar (i.e. gasto per cápita del hogar) y su dinámica temporal propia; (iii) Análisis de Componentes Principales (PCA, por sus siglas en inglés); y (iv) determinación de los puntos de corte en base a las metodologías especificadas en el capítulo III del presente documento.

##### A. Selección de Variables mediante Análisis de Correlación y Dinámica Temporal

Puesto que se pretende captar la dinámica de las interrelaciones entre diversas variables socioeconómicas del hogar con la condición de pobreza o el nivel de consumo per cápita del mismo, se construye una base de datos apilada que contiene información de varios cortes de la Encuesta Nacional de Hogares (ENAHO). En específico, se considera un periodo de ocho años que comprende desde el año 2007 hasta el año 2014. Esto nos permite, además de enfatizar en el

carácter temporal de las mencionadas asociaciones, captar diferentes periodos de la economía peruana, tanto de expansión como de contracción económica, develando relaciones estructurales que exceden a las que se encontrarían en coyunturas específicas.

La decisión metodológica asimismo permite una inferencia más precisa a un menor nivel de desagregación, ampliando las posibilidades de análisis. En este caso, se establece que la estimación de los modelos de IFH se realice a nivel de 15 conglomerados geográficos, compuestos por las combinaciones de los ocho dominios identificables en la ENAHO<sup>7</sup> y la condición de ruralidad (ámbito urbano o rural), la que difiere a nivel de centro poblado<sup>8</sup>.

Se realiza una inspección de las variables que coinciden entre la Ficha Socioeconómica Única y la Encuesta Nacional de Hogares. Para ello, se realiza la desagregación de la información para analizar el comportamiento de variables que tengan una relación significativa con la pobreza monetaria y con el gasto per cápita<sup>9</sup>. El análisis de Componentes Principales requiere que las variables empleadas tengan una alta correlación con la variable de interés. Dichos cortes se realizaron por años, dominios, estratos, área, conglomerados, deciles de gasto per cápita por departamento y por conglomerados.

Una vez que estas variables se consideraban fuertemente correlacionadas con el gasto y con la pobreza monetaria, eran incluidos en la siguiente etapa. La naturaleza del método de PCA, no permitía un único método de selección de indicadores pues su naturaleza es capturar la mayor varianza conjunta posible de las variables incluidas. Así que muchas veces variables con una gran correlación con la pobreza o el gasto no eran incluidos finalmente en el modelo pues su varianza ya era fuertemente explicada por otras variables.

Si bien no hay un método claro para definir que variables deben o no incluirse, la elección se determinó, en mayor medida por los siguientes criterios: (i) capturar la mayor cantidad posible de "dimensiones" importantes de la pobreza, tales como el acceso a bienes, servicios públicos y educación; (ii) tener una correlación estable en el tiempo, es decir, que no se espere que pierda o gane un poder predictivo con cambios en el tiempo; (iii) lograr niveles altos de áreas bajo las curvas ROC<sup>10</sup> del primer componente, el criterio más utilizado en la evaluación de Proxy Mean Tests; (iv) obtener un factor (producto de PCA) que se distribuya de forma normal o, que parezca ser así; y, (v) En caso de tenerse más de un factor para un conglomerado, unificarlos de manera que se maximice el poder predictivo conjunto.<sup>11</sup>

En el gráfico 3 se presenta un ejemplo de una de las variables incluidas en los modelos PCA realizados para uno de los ámbitos. En este caso se observa como la variable años de educación del jefe del hogar está claramente vinculada con el gasto per cápita por decil. Así, esta variable fue incluida en el modelo predictor del conglomerado de Lima Metropolitana pues muestra tener una correlación elevada, además de que esta es muy estable en el tiempo. En el Anexo A se presentan las estadísticas pertinentes en relación a las variables de interés.

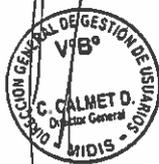
<sup>7</sup> Los dominios en la ENAHO son una combinación de la región natural (Costa, Sierra o Selva) donde se encuentra el centro poblado y el departamento en el que se encuentra. Así, la zona norte se compone por los departamentos: Amazonas, Cajamarca, La Libertad, Lambayeque, Piura y Tumbes. La zona centro se compone: Ancash, Ayacucho, Huancavelica, Huánuco, Ica, Junín, Lima (sin Lima Metropolitana) y Pasco. La zona sur se compone por: Apurímac, Arequipa, Cusco, Moquegua, Puno y Tacna.

<sup>8</sup> Se considera a un centro poblado como rural si tiene menos de 400 viviendas.

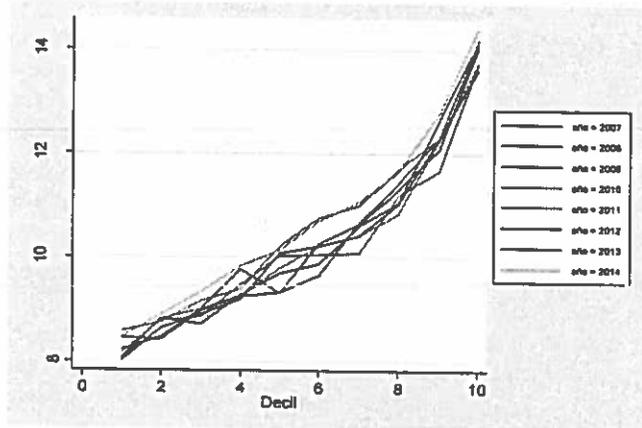
<sup>9</sup> Con fines de homogenización del análisis, el gasto per cápita se deflactó a nivel geográfico (a soles de Lima Metropolitana) y temporalmente (al 2014).

<sup>10</sup> Las curvas ROC (Receiving Operating Curves, en inglés) presentan la capacidad predictiva del índice (en este caso el componente principal obtenido por el PCA).

<sup>11</sup> Se elegirían umbrales para cada IFH y se comparará si se obtiene mejor predictivo combinando ambos de manera complementaria (ambos tienen que decir que es pobre para que sea considerado como tal) o sustitutoria (basta con uno que indique que es pobre para ser considerado como tal).



**Gráfico 3. Educación promedio del jefe del hogar por decil de gasto per cápita en Lima Metropolitana para los años 2007-2014**

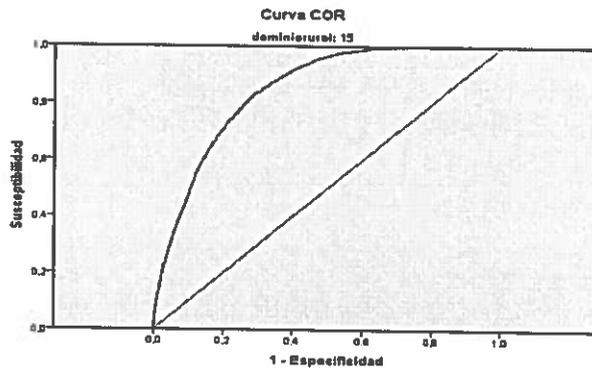


Fuente: ENAHO 2007-2014

Elaboración: DO-DGGU

Entonces, si bien esto nos da criterios para incluir o excluir alguna variable del modelo, el criterio siguiente es más importante pues lo que se busca es que se tenga un factor principal (predicho por PCA) que logre una gran capacidad predictiva de la pobreza monetaria. El criterio al cual se busca aproximar. El gráfico 4 presenta un ejemplo de curva ROC de uno de los factores predichos para Lima Metropolitana, mientras el gráfico 5 presenta la distribución de uno de los factores predicho. Para el criterio de la curva ROC, se debe buscar que esta cubra la mayor parte del gráfico, siendo en el caso hipotético de perfecta predicción que lo cubra totalmente y en el peor caso que se la recta de 45°, que indicaría una capacidad de predicción similar al azar.

**Gráfico 4. Curva ROC de un factor predicho por PCA de su capacidad predictiva de la pobreza monetaria en Lima Metropolitana para los años 2007-2014**



Fuente: ENAHO 2007-2014 Elaboración: DO-DGGU

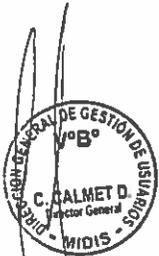
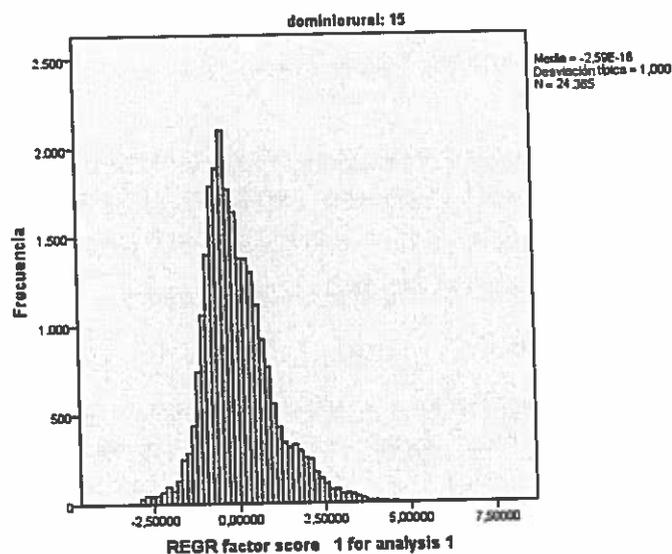


Gráfico 5. Histograma de un factor predicho por PCA en Lima Metropolitana para los años 2007-2014



ENAH0 2007-2014

Elaboración: DO-DGGU

### B. Análisis de Componentes Principales - PCA

El IFH, al igual que en la Metodología MEF 2010, se estima usando el método de componentes principales (PCA, por las siglas de Principal Component Analysis), salvo que en este caso no se realiza el escalamiento óptimo (OS, por las siglas de Optimal Scaling) por medio de la opción de OS, sino a través de la estandarización de la variable en relación a la media y desviación estándar de la variable en el periodo 2007-2014.

El PCA es un método que busca reducir la dimensionalidad de un conjunto de variables aleatorias conservando la mayor parte de la información (i.e. de la variabilidad). De esta forma, para un vector de  $p$  variables aleatorias  $x$ , con matriz de varianzas y covarianzas  $V[x]$ , este proceso requiere hallar las direcciones de mayor varianza de las combinaciones lineales de las variables aleatorias. En rigor, esto corresponde a encontrar el conjunto de  $K \leq p$  vectores ortonormales  $a_1, \dots, a_k$  que resuelven el conjunto de problemas de maximización:

$$\begin{aligned}
 a_1 &= \arg \max_{a: |a|=1} V[a'x] \\
 a_2 &= \arg \max_{\substack{a: |a|=1 \\ a \perp a_1}} V[a'x] \\
 &\dots \\
 a_k &= \arg \max_{\substack{a: |a|=1 \\ a \perp a_1, \dots, a \perp a_{k-1}}} V[a'x]
 \end{aligned}
 \tag{1}$$

De esta manera, cada vector de ponderadores  $a_k$  es ortogonal a los vectores previamente hallados. En ese sentido, la información recogida por cada vector  $a_k$  es única, pues corresponde a la



### Metodología para la Determinación de la Clasificación Socioeconómica

variabilidad del conjunto de variables  $x$  en una dirección en particular, que es ortogonal a la variabilidad en otras direcciones.

El PCA resulta útil porque permite construir variables que resumen la información contenida en un conjunto mayor de variables. Así, los vectores  $a_k$  obtenidos de (1) permiten construir variables a las que se llaman “componentes”, para un vector dado  $x = (x_1, x_2, \dots, x_p)$  de  $p$  variables aleatorias. De esta forma, el  $k$ -ésimo componente es definido como:

$$y_k = a_k x \equiv a_{1k} x_1 + a_{2k} x_2 + \dots + a_{pk} x_p \quad (2),$$

En este contexto, el primer componente,  $y_1$  explica una fracción más importante de  $V[x]$  que los demás componentes siguientes; el segundo componente  $y_2$  explica la segunda fracción más importante de  $V[x]$ , y así sucesivamente. Esta característica facilita la interpretación directa de los resultados del PCA en los *proxy means test* (PMT), pues permite combinar información de diferentes variables que aproximen la situación de bienestar de los hogares en la construcción de un único índice, usando como ponderadores los elementos de  $a_k$ .

En la práctica, esto supone elegir el número de componentes en que se quiere resumir la información de las  $p$  variables. Por simplicidad, lo usual es conservar solo el primer componente  $y_1$  (Bin Zahur, 2009), y usar como ponderadores a los elementos del vector  $a_1$ . Dicho índice luego es considerado como un indicador suficiente de la calidad de vida del hogar.

Un problema frecuente y no menor en la implementación del PCA es que las variables elegidas por los PMT no son usualmente cardinales y, por ende, no son continuas. Este es el caso de los indicadores de calidad de vida, que si bien son usualmente ordinales, resultan muy confiables pues contribuyen con información difícilmente manipulable por el hogar. En consecuencia, se hace necesario transformar estas variables para poder aplicar el PCA.

La transformación de variables ordinales en cardinales se puede hacer de distintas formas; la adoptada es la dicotomización de las categorías de las variables y su estandarización en el periodo 2007-2014.

### C. Determinación de Puntos de Cortes para la CSE

Para la determinación de los puntos de corte se emplean las metodologías mencionadas previamente en el capítulo III, las cuales son: i) Minimización del Error de Focalización Ponderado; y, ii) Umbral de Exclusión.

La Minimización del error de focalización reporta el umbral que minimiza la función de error total del instrumento  $p$ , definida como la suma del error de inclusión ( $EI_p$ ) y error de exclusión ( $EE_p$ ). El  $EE$  decrece y  $EI$  crece a medida que el umbral ( $U$ ) aumenta. En este sentido, se puede encontrar, bajo algunas condiciones, el umbral que minimice la fórmula:

$$ET(U) = \emptyset EE_p(U) + EI_p(U) \quad (3)$$

El cual deberá cumplir siguiente condición matemática:

$$\frac{\partial ET(U)}{\partial U} = \frac{\emptyset \partial EE_p(U)}{\partial U} + \frac{\partial EI_p(U)}{\partial U} = 0 \quad (3')$$

El parámetro  $\emptyset$  es una constante que representa el costo en términos de error de inclusión que el hacedor de política se encuentra dispuesto a incurrir por la reducción en un 1 punto porcentual del error de exclusión. De esta forma, matemáticamente se podría presentar como una minimización de la suma de los errores pero de manera ponderada. Para fines del ejercicio, se emplea la ponderación de dos (2), siendo la fórmula a estimar la siguiente:

$$ET(U) = 2 * EE_p(U) + EI_p(U) \quad (4)$$

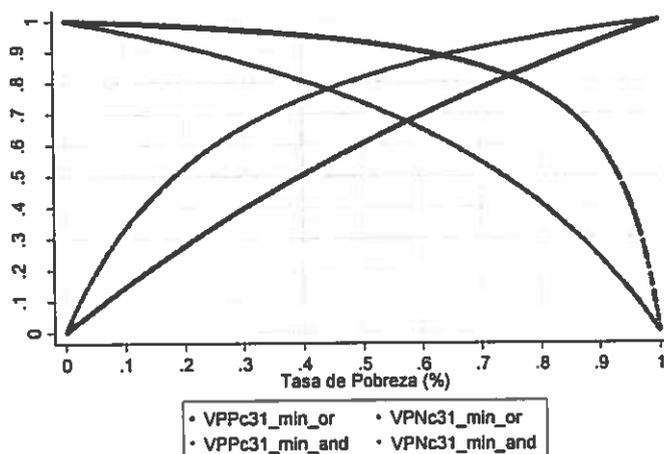
Por su parte, el Umbral de Exclusión reporta el mínimo punto de corte ( $U$ ) que reporte el error de exclusión de un 2%; es decir en donde:

$$EE_p(U) = 2\% \quad (5)$$

Seguidamente, en función a la prevalencia de la pobreza en el conglomerado en mención, se escogen los umbrales de los índices que conforman el IFH, estos son aplicados de forma paralela bajo un criterio de unión (OR) o de intersección (AND). La decisión final sobre la determinación del punto de corte, dependerá del aporte marginal en cuanto al valor predictivo; siendo deseable mejoras en el VPN para conglomerados con alta incidencia de la pobreza y mejoras en el VPP para conglomerados de baja incidencia.

Para ejemplificar observe el siguiente Gráfico 6. Si la siguiente gráfica proviene de un conglomerado con una tasa de pobreza de 70%; priorizare el VPN, a fin de mejorar la precisión de mi estimación de No pobreza. Por tanto, deberé escoger la combinación "OR", puesto que alrededor del 70% el VPN sube de 40% a 80%. Es decir, una vez obtenido el resultado de No pobre, la probabilidad de que ese individuo sea efectivamente No pobre se duplica. Esta lógica guiará la selección de umbrales.

**Gráfico 6. Valores Predictivos Positivos y Negativos de dos pruebas aplicadas en paralelo con cortes estimados minimizando el error de focalización**



Fuente: ENAHO (2014)

Fuente: ENAHO 2007-2014  
Elaboración: DO-DGGU

#### D. Criterios de Evaluación de la Capacidad Predictiva

Para evaluar la capacidad predictiva del algoritmo se siguen distintos criterios que son considerados estándar por la literatura especializada: (i) Sensibilidad, porcentaje de pobres asignados correctamente; (ii) especificidad, porcentaje de no pobres asignados correctamente; (iii) valor predictivo positivo, porcentaje de los asignados como pobres que efectivamente lo son; (iv) valor predictivo negativo, porcentaje de los asignados como no pobres que efectivamente no son pobres; y, finalmente, el área debajo de la curva ROC, que es un indicador de la capacidad predictiva del IFH de manera conjunta.

Algebraicamente estos indicadores se pueden expresar en las siguientes ecuaciones:

## Metodología para la Determinación de la Clasificación Socioeconómica

*Sensibilidad:*

$$\text{Sensibilidad} = 1 - EE_p = 1 - \frac{p_{10}}{NP} \quad (6),$$

Donde  $p_{10}$  representa el número de individuos pobres monetarios que son clasificados como no pobres por el algoritmo  $pp$ . La sensibilidad se puede entender como la contraparte del error de exclusión, indica cuantos de los pobres son efectivamente asignados como tal.

En este sentido, un alto valor de sensibilidad sugiere que el algoritmo no "excluye" a la población efectivamente pobre.

*Especificidad:*

$$\text{Especificidad} = 1 - EI_p = 1 - \frac{p_{01}}{N^{np}} \quad (7),$$

Donde  $p_{10}$  es el número de individuos no pobres monetarios que son clasificados como pobres por el algoritmo  $p$ , y  $N^{np}$ , la población no pobre de acuerdo con el INEI. La especificidad puede considerarse como la contraparte del error de inclusión, así un valor alto de la misma nos indica que tenemos éxito en asignar a los individuos no pobres, tal condición.

*Valor Predictivo Positivo:*

$$VPP_p = \frac{p_{11}}{p_{10} + p_{11}} \quad (8),$$

Donde  $p_{10}$  representa el número de individuos pobres monetarios que son clasificados como no pobres por el algoritmo mientras que  $p_{11}$  representa a los pobres monetarios clasificados como tal por el algoritmo. En tal sentido se puede decir que es la probabilidad de que una vez que has sido asignado como pobre, lo seas efectivamente.

*Valor Predictivo Negativo:*

$$VPN_p = \frac{p_{00}}{p_{01} + p_{00}} \quad (9),$$

Donde  $p_{01}$  representa el número de individuos no pobres monetarios que son clasificados como pobres por el algoritmo mientras que  $p_{00}$  representa a los no pobres monetarios clasificados como tal por el algoritmo. En tal sentido se puede decir que es la probabilidad de que una vez que has sido asignado como no pobre, lo seas efectivamente.

*Curvas ROC*

Este análisis sigue el análisis propuesto en el trabajo de Wodon (1997)<sup>12</sup>, quien emplea las curvas ROC para comparar el desempeño de un conjunto de instrumentos de focalización para identificar a los pobres.

Para evaluar la capacidad discriminatoria entre pobres y no pobres, se utilizan curvas ROC, las mismas que se basan en la sensibilidad (probabilidad de clasificar correctamente a un individuo cuyo estado real sea el definido como positivo respecto a la característica socioeconómica evaluada) y especificidad (probabilidad de clasificar correctamente a un individuo cuyo estado real sea el definido como negativo).

Este tipo de análisis resulta aplicable cuando se dispone de una variable ordinal que intenta predecir una condición y una variable dicotómica igual a uno si la condición efectivamente se registra (estado real). En este contexto, se requiere un nivel de corte de tal modo que, si la variable ordinal supera el umbral, la predicción es que la condición es verdadera. El procedimiento implica asumir todos los valores de corte posibles y, en cada caso, calcular dos indicadores denominados sensibilidad y especificidad. Se define sensibilidad como el número de casos en los que el Índice

<sup>12</sup> Wodon, Q (1997). Targeting the Poor Using ROC Curves. World Development, 25(12), 2083-2092.

predijo que la condición era cierta entre el número casos en los que la condición efectivamente se registró. Asimismo, se define especificidad como el número de casos en los que la condición no se cumple según el modelo, dividido entre el número de casos en los que la condición no se cumple en la realidad. El resultado de esta etapa es una tabla donde, para cada posible nivel de corte, existe una sensibilidad y una especificidad.

Posteriormente, se obtiene la curva ROC como resultado de la unión de los pares ordenados (1-especificidad, sensibilidad). La intuición detrás consiste en que si la variable ordinal predice perfectamente a la variable dicotómica, el área debajo de la curva ROC es igual a 1. Si la variable ordinal predice a la variable dicotómica de manera completamente imperfecta, el área debajo de la curva ROC es igual a 0. Si la variable ordinal predice a la variable dicotómica con una probabilidad igual a 0.5, la curva ROC es una recta que determina un área igual a 0.5. En este contexto, el área debajo de la curva ROC es un indicador del éxito de la variable ordinal prediciendo la variable dicotómica y se espera que por lo menor sea mayor a 0.5.

## V. Índice de Focalización de Hogares (Algoritmo) y Capacidad Predictiva

### A. Selección de variables mediante análisis de correlación y dinámica temporal

Los conglomerados considerados son los siguientes: Costa Norte Rural, Costa Norte Urbana, Costa Centro Rural, Costa Centro Urbana, Costa Sur Rural, Costa Sur Urbana, Sierra Norte Rural, Sierra Norte Urbana, Sierra Centro Rural, Sierra Centro Urbana, Sierra Sur Rural, Sierra Sur Urbana, Selva Rural, Selva Urbana, y Lima Metropolitana. Así la muestra final se presenta por conglomerado en el Tabla 1. Cabe precisar que anualmente se recolecta información de al menos 20,000 hogares.

Tabla 1. Muestra Apilada por Conglomerado Geográfico  
ENAH0 2007-2014

Dominio Geográfico	U/R	Muestra
Costa Norte	Rural	5,476
	Urbana	21,262
Costa Centro	Rural	2,949
	Urbana	12,723
Costa Sur	Rural	1,818
	Urbana	8,894
Sierra Norte	Rural	9,592
	Urbana	2,849
Sierra Centro	Rural	22,610
	Urbana	15,111
Sierra Sur	Rural	16,141
	Urbana	13,244
Selva	Rural	18,565
	Urbana	22,549
Lima Metropolitana		24,365
TOTAL		198,148

Fuente: ENAH0 2007-2014  
Elaboración: DO-DGGU

Para la construcción del IFH, se han evaluado alrededor de 115 variables agrupadas en cinco categorías: (i) bienes del hogar, (ii) servicios básicos del hogar, (iii) calidad de la vivienda, (iv) demográficas y (v) capital humano. En el Anexo A se presentan las variables que cumplieron con los criterios temporal y de correlación con la medida oficial de pobreza. Las variables que



*Metodología para la Determinación de la Clasificación Socioeconómica*

encuentran mayor correlación con la pobreza son los activos del hogar y las características de la vivienda.

**B. Análisis de Componentes Principales - PCA**

Luego de excluir las variables que no detentan una correlación fuerte con la pobreza, según el análisis de deciles y las que dicha asociación era inestable en el tiempo, se construye el IFH para quince ámbitos con las variables restantes. En la mayoría de casos con las variables seleccionadas no fue posible retener solo un componente que explicará más del 50% de la varianza. Inclusive cuando esto fue posible, existía un grupo amplio de variables que también aportaban información que, sin embargo, según el PCA, no se encontraba en el mismo componente. Es decir, brindaban información que explicaba gran parte de la varianza pero que no se encontraba necesariamente en la misma dirección de las variables excluidas. Por tanto, la integración dentro de un solo componente de todas las variables implicaba pérdidas de eficiencia (menor varianza explicada).

En este contexto, se optó por correr dos PCA de forma separada cuando fue necesario y retener el primer componente de los grandes grupos de variables- denominados Factores. La regularidad empírica evidencia la conformación de dos grupos: por un lado asociados a los servicios básicos de la vivienda (grupo ii), calidad de la vivienda (grupo iv), demográficas (grupo iv) y capital humano (grupo v) y por otro, las variables relacionadas a los activos del hogar (grupo i). En todos los casos, la varianza explicada por el primer componente supera el estándar de 80% y presenta áreas debajo de la curva ROC de más de 70%; lo que implica una buena medida de ajuste a la data y una capacidad predictiva de la pobreza satisfactoria.

En general, las puntuaciones registradas tienen el orden que se esperaba; es decir, los valores que aportan positivamente son los que corresponden a las categorías que están asociadas a un mayor nivel de precariedad; mientras que, las variables que aportan negativamente se asocian a un menor nivel de precariedad (Tabla 2 y 3).

Vale precisar que para el caso del ámbito rural, todos los modelos ajustaron bien los datos en un solo factor por conglomerado; salvo el caso de la costa norte y centro del ámbito rural, en los que se requirió la conformación de modelos con dos factores.

**Tabla 2 Matriz de Coeficientes de las Variables del 1° y 2° Factor por Conglomerado del Área Urbana- IFH 2015**

	Costa						Sierra						Selva		Lima	
	Norte		Centro		Sur		Norte		Centro		Sur		F1	F2	F1	F2
	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
Ratio de hacinamiento	0.20		0.24		0.26		0.60		0.30		0.13		0.25		0.37	
Presencia de piso de tierra	0.33		0.36		0.35		0.14		0.35		0.29		0.33		0.31	
Usa combustible sólido para cocinar	0.28		0.24		0.21		0.14				0.27		0.34		0.20	
Presencia de techo distinto a concreto	0.29		0.33		0.35		0.09		0.31		0.26		0.23		0.35	
Hogar sin ningún servicio Púb. Conect.	0.25								0.27		0.18		0.24			
Hogar con los tres servicios Púb. Conect.			-0.31		-0.30											
Hogar con jefe asegurado			-0.23													
Hogar con jefe con lengua indígena					0.17						0.21					
Máximo años de educación en el hogar	-0.27		-0.16				-0.15				-0.24		-0.30			
Años de educación del jefe de hogar															-0.14	
N° de bienes en el hogar		-0.21		-0.26		-0.27	-0.21			-0.30	-0.26		-0.26		-0.28	
N° de bienes durables en el hogar		-0.19		-0.24		-0.26	-0.20			-0.28	-0.25		-0.25		-0.26	
N° de electrodomésticos en el hogar		-0.21		-0.23		-0.25	-0.19			-0.27	-0.24		-0.24		-0.25	
N° de bienes de telecomunicaciones		-0.16		-0.21		-0.22	-0.17			-0.24	-0.22		-0.22		-0.23	
N° de bienes per cápita		-0.14		-0.13		-0.17		-0.60	-0.37		-0.17		-0.18			-0.34
Tenencia de licuadora en el hogar		-0.16														
Tenencia de plancha eléct. en el hogar		-0.15														

Fuente: ENAHO 2007-2014  
Elaboración: DO-DGGU

Tabla 3. Matriz de coeficientes de las variables escogidas del 1º y 2º factor por conglomerado del área rural- IFH 2015

	Costa						Sierra			Selva
	Norte		Centro		Sur	Norte	Centro	Sur		
	F1	F2	F1	F2	F1	F1	F1	F1		
Nº de electrodomésticos en el hogar	-0.19		-0.29			-0.23	-0.21	-0.22	-0.21	
Nº de bienes en el hogar	-0.19		-0.31				-0.22	-0.22	-0.23	
Nº de bienes durables en el hogar	-0.19		-0.30			-0.25	-0.21	-0.22	-0.22	
Nº de bienes per cápita	-0.15		-0.23		-0.54	-0.24	-0.23	-0.24	-0.20	
Tenencia de refrigeradora/ congeladora en el hogar	-0.16									
Tenencia de cocina a gas en el hogar	-0.15									
Tenencia de licuadora en el hogar	-0.16									
Usa combustible sólido para cocinar		0.41	0.42	0.43		0.17	0.16		0.14	
Usa gas para cocinar		-0.41								
Presencia de piso de tierra		0.27	0.48			0.19	0.16	0.17		
Presencia de techo distinto a concreto		0.12	0.39							
Ratio de hacinamiento		0.15	0.26	0.41		0.08	0.14	0.15	0.09	
Máximo años de educación en el hogar							-0.06	-0.06		
Hogar conectado a desagüe por red púb. dentro						-0.16		-0.14		
Jefe de hogar con secundaria incompleta						0.14				
Hogar conectado a energía eléc. por red públ.									-0.16	

Fuente: ENAHO 2007-2014

Elaboración: DO-DGGU

### C. Determinación de Puntos de Cortes para la CSE

A partir del ejercicio propuesto, se identifican 15 conglomerados de zonas geográficas. Para definir umbrales se siguen los siguientes pasos. Primero, se calculan los umbrales a partir de la metodología de minimización del error total de focalización (la versión pondera y sin ponderar) y la del umbral del error de exclusión del 2%. Seguidamente, en función a la prevalencia de la pobreza en el conglomerado en mención, se escogen los umbrales de los índices que conforman el IFH, estos son aplicados de forma paralela bajo un criterio de unión o de intersección. La decisión final sobre la determinación del punto de corte, dependerá del aporte marginal en cuanto al valor predictivo; siendo deseable mejoras en el VPN para conglomerados con alta incidencia de la pobreza y mejoras en el VPP para conglomerados de baja incidencia.

Tabla 4. Umbrales de Corte de los Factores por Conglomerado- IFH 2015

Dominio Geográfico	U/R	Factor 1	Factor 2
Costa Norte	Urbano	-0.22	0.01
	Rural	-1.10	-1.17
Costa Centro	Urbano	0.03	0.15
	Rural	0.18	0.33
Costa Sur	Urbano	0.13	-0.21
	Rural	0.80	
Sierra Norte	Urbano	0.30	0.38
	Rural	-1.34	
Sierra Centro	Urbano	0.17	0.18
	Rural	-0.68	
Sierra Sur	Urbano	0.60	0.09
	Rural	-0.37	
Selva	Urbano	0.21	0.26
	Rural	-0.78	
Lima Metropolitana		-0.18	-0.04

Fuente: ENAHO 2007-2014

Elaboración: DO-DGGU



D. Evaluación de la Capacidad Predictiva

La tabla 5 presenta el área debajo de la curva ROC de cada uno de los factores (o primeras componentes) por conglomerado. La buena capacidad de predicción de los puntajes calculados y la condición de pobreza permite usarlo como instrumento eficiente para identificar a la población en situación de carencia, con la ventaja de ser construido con las variables de la FSU con un enfoque que recogería diversas dimensiones de la pobreza.

Tabla 5. Área debajo de la Curva ROC por Factor y Conglomerado Geográfico ENAHO 2007-2014

Dominio Geográfico	U/R	Factor 1	Factor 2
Costa Norte	Urbano	81.0%	80.5%
	Rural	78.5%	76.2%
Costa Centro	Urbano	81.8%	81.3%
	Rural	77.6%	76.3%
Costa Sur	Urbano	80.1%	83.1%
	Rural	83.5%	
Sierra Norte	Urbano	87.6%	84.5%
	Rural	76.2%	
Sierra Centro	Urbano	82.7%	85.0%
	Rural	77.5%	
Sierra Sur	Urbano	84.9%	87.0%
	Rural	78.4%	
Selva	Urbano	83.0%	82.9%
	Rural	76.8%	
Lima Metropolitana		80.9%	84.6%

Fuente: ENAHO 2007-2014  
Elaboración: DO-DGGU

Para la selección de umbrales se tomó en cuenta que estos instrumentos son usados como línea de detección de pobreza del hogar solicitante, por lo cual se requiere altos valores de sensibilidad que mejoren el rendimiento de la focalización de hogares. Por ello no solo es necesario evaluar los valores de error de inclusión / exclusión (y sus contrapartes sensibilidad / especificidad) sino que se evaluó el rendimiento predictivo para las frecuencias de pobreza que se esperaba se encuentre en campo, tratando que la predicción negativa no baje del 90%. En relación a los puntos de corte empleados, teniendo en consideración la tasa de incidencia de la pobreza específica a los conglomerados, se obtienen valores predictivos negativos por encima del 80%.



Tabla 6. Sensibilidad, Especificidad y Valores Predictivos Positivos y Negativos- IFH 2015

Dominio Geográfico	U/R	Tasa de Pobreza	Sensibilidad	Especificidad	Valor predictivo positivo (Máximo)*	Valor predictivo negativo (Máximo)
Costa Norte	Urbano	20.43%	69%	87%	58% (82%)	92% (98%)
	Rural	38.40%	94%	45%	52% (66%)	92% (93%)
Costa Centro	Urbano	9.47%	70%	89%	40% (69%)	97% (98%)
	Rural	9.41%	43%	94%	42% (56%)	94% (94%)
Costa Sur	Urbano	10.26%	70%	84%	34% (63%)	96% (98%)
	Rural	10.68%	96%	87%	47% (61%)	99% (99%)
Sierra Norte	Urbano	26.54%	70%	91%	73% (90%)	89% (98%) †
	Rural	65.24%	98%	21%	70% (80%)	85% (88%) †
Sierra Centro	Urbano	21.16%	70%	90%	65% (87%)	92% (98%)
	Rural	48.30%	95%	38%	59% (72%)	89% (88%) †
Sierra Sur	Urbano	13.12%	66%	95%	68% (87%)	95% (98%)
	Rural	39.84%	92%	40%	51% (64%)	89% (88%) †
Selva	Urbano	22.61%	62%	91%	67% (87%)	89% (98%) †
	Rural	41.52%	94%	38%	52% (66%)	90% (90%)
Lima Metropolitana		11.78%	79%	87%	44% (70%)	97% (97%)

Fuente: ENAHO 2007-2014

Elaboración: DO-DGGU

\*El valor predictivo máximo se presenta entre paréntesis y se alcanza cuando el hogar puede ser evaluado con información disponible para ser contrastada con los cortes de ingresos y de consumo de energía eléctrica además del índice.

†Se decidió mantener estos modelos, aunque no alcanzan valores predictivos negativos mayores de 90%, puesto que mostraron mejores valores predictivos positivos.

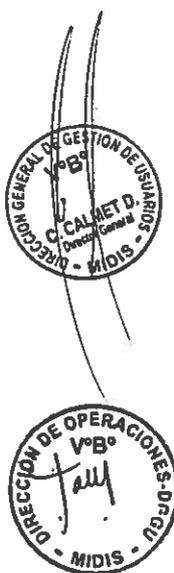
## VI. Conclusiones

La operatividad del SISFOH está a cargo de la UCF, la cual se encuentra bajo el ámbito de la DGGU. Asimismo, la UCF tiene como función determinar y certificar la CSE de los potenciales usuarios ante los agentes responsables de la administración de los programas sociales y de subsidios del Estado.

La UCF determina la CSE de las personas y sus hogares sobre la base de la "Metodología de Cálculo del Índice de Focalización de Hogares" aprobada mediante Resolución Ministerial N° 320-2010-PCM (Metodología MEF-2010), la cual se elaboró con información de la ENAHO 2009. Asimismo, se advierte que desde el pasado 23 de abril del 2015, se cuenta con información de la ENAHO 2014, motivo por el cual corresponde la revisión de los parámetros y la evaluación de los instrumentos empleados.

Como resultado, de dicha revisión y evaluación, corresponde aprobar la Metodología para la Determinación de la Clasificación Socioeconómica, que reemplace a la Metodología MEF-2010.

Con esta nueva metodología se propone la incorporación de cambios conceptuales en relación a la medición de los recursos económicos (familiares y no individuales), considerando la incorporación de las características del hogar (tamaño del grupo familiar) para su cálculo y la dinámica territorial diferenciada según ámbito (Lima Metropolitana, Resto Urbano y Rural) para la evaluación de los umbrales de corte, en función de las diferencias en el dinamismo de la actividad económica, provisión de servicios públicos y desarrollo de los mercados que se reflejan en la capacidad de consumo y gasto de los hogares y en la posibilidad que tienen de aprovechar las oportunidades económicas.



### Metodología para la Determinación de la Clasificación Socioeconómica

Asimismo, se extienden las metodologías empleadas para el cálculo de umbrales a partir de la optimización de diferentes funciones objetivo (minimización de errores de focalización ponderados y el mínimo error de exclusión de 2% y las herramientas de análisis, asociando a cada punto de corte los valores predictivos, definidos como la probabilidad de que el hogar sea realmente un hogar pobre (o no) ante un resultado concreto de la metodología. Esta mejora se incorpora en la evaluación de la secuencia de pasos dentro de la metodología, seleccionando los arreglos que presentan un mayor aporte en la predicción de la pobreza, medido como el cambio marginal de los valores predictivos.

Adicionalmente, atendiendo a los diferentes cambios en los procesos y a las funciones del SISFOH, se definen mejoras en la terminología, planteando el cambio de "Procedimiento de Determinación de Elegibilidad" a "Metodología para la Determinación de la Clasificación Socioeconómica" y la incorporación de definiciones operativas para dicho proceso. Se propone un Glosario de Términos que se adjunta en el Anexo B del presente documento.

La metodología propuesta se resume en:

- (i) Evaluación del Nivel de Ingresos: Se contrasta el ingreso neto por planilla pública o privada per cápita con el punto de corte de S/.383 Nuevos Soles para Lima Metropolitana, S/. 337 Nuevos Soles para el resto urbano y S/. 320 Nuevos Soles para el ámbito rural. Aquellos hogares que sobrepasen estos montos serán considerados como no pobres y no seguirán al paso (ii).
- (ii) Evaluación del gasto en consumo de agua y electricidad: Se compara el consumo de electricidad con el punto de corte de S/. 18, S/ 10.1 y S/ 9.2 para los ámbitos de Lima Metropolitana, resto urbano y rural respectivamente. Los hogares con un consumo per cápita superior son considerados como no pobres y no seguirán al paso (iii).
- (iii) Evaluación de la calidad de vida del hogar: El IFH-2015 será aplicado a los hogares que se hayan encontrado por debajo de los umbrales anteriores o que no hayan tenido información previa. Este resultado, interpretado como No Pobre, Pobre o Pobre Extremo según los puntos de cortes emitidos por la Dirección General de Gestión de Usuarios, será publicado en el Padrón General de Hogares.



VII. Anexos

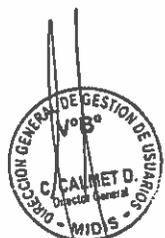
ANEXO A1. Relación de las variables escogidas para primer factor y su relación con los deciles de gasto per cápita por conglomerado para la base apilada en ENAHO 2007-2014

		Decil del gasto per cápita en el hogar									
		1	2	3	4	5	6	7	8	9	10
Costa Norte Urbana	Ratio de hacinamiento	2.26	1.94	1.79	1.62	1.51	1.35	1.23	1.06	0.93	0.69
	Presencia de piso de tierra	0.68	0.54	0.48	0.43	0.34	0.27	0.20	0.13	0.10	0.04
	Usa combustible sólido para cocinar	0.55	0.41	0.38	0.31	0.23	0.18	0.13	0.08	0.05	0.02
	Presencia de techo distinto a concreto	0.93	0.87	0.84	0.81	0.74	0.70	0.62	0.57	0.46	0.30
	Hogar sin ningún servicio público conectado	0.48	0.34	0.29	0.22	0.19	0.16	0.11	0.08	0.06	0.04
	Años de educación máximo en el hogar	9.24	10.15	10.27	10.75	11.13	11.44	12.02	12.47	12.93	13.94
Costa Norte Rural	Ratio de hacinamiento	2.04	1.72	1.54	1.33	1.17	1.10	0.97	0.83	0.78	0.64
	Presencia de piso de tierra	0.85	0.79	0.72	0.70	0.58	0.56	0.52	0.47	0.40	0.28
	Usa combustible sólido para cocinar	0.92	0.87	0.79	0.75	0.68	0.59	0.53	0.51	0.43	0.37
	Presencia de techo distinto a concreto	0.99	0.99	0.99	0.98	0.98	0.96	0.97	0.97	0.95	0.90
	Usa GLP para cocinar	0.08	0.13	0.21	0.25	0.31	0.41	0.47	0.48	0.56	0.61
Costa Centro Urbana	Ratio de hacinamiento	2.59	1.81	1.98	1.68	1.38	1.26	1.26	1.05	0.95	0.63
	Presencia de piso de tierra	0.46	0.37	0.29	0.21	0.20	0.15	0.11	0.10	0.06	0.03
	Usa combustible sólido para cocinar	0.21	0.19	0.10	0.07	0.09	0.02	0.03	0.03	0.02	0.01
	Presencia de techo distinto a concreto	0.86	0.69	0.77	0.64	0.62	0.54	0.51	0.39	0.36	0.22
	Años de educación máximo en el hogar	10.20	11.00	11.95	11.75	11.16	11.97	12.40	12.74	12.80	13.59
	Hogar sin ningún servicio público conectado	2.13	2.28	2.47	2.60	2.60	2.59	2.66	2.75	2.84	2.96
Costa Centro Rural	Presencia de piso de tierra	0.57	0.41	0.43	0.39	0.32	0.31	0.30	0.20	0.19	0.17
	Presencia de techo distinto a concreto	0.93	0.86	0.88	0.83	0.84	0.81	0.77	0.77	0.76	0.67
	Usa combustible sólido para cocinar	0.63	0.49	0.46	0.38	0.34	0.35	0.34	0.20	0.24	0.14
	Ratio de hacinamiento	1.84	1.58	1.39	1.29	1.20	1.07	1.02	0.86	0.71	0.52
Costa Sur Urbana	Ratio de hacinamiento	2.26	1.99	1.74	1.57	1.40	1.21	1.10	0.94	0.78	0.58
	Presencia de piso de tierra	0.34	0.24	0.19	0.15	0.10	0.08	0.07	0.05	0.03	0.01
	Usa combustible sólido para cocinar	0.27	0.17	0.12	0.08	0.06	0.04	0.04	0.03	0.01	0.00
	Presencia de techo distinto a concreto	0.60	0.51	0.41	0.38	0.33	0.28	0.25	0.21	0.15	0.11
	Hogar con los tres servicios públicos conectados en el hogar	2.61	2.68	2.76	2.83	2.83	2.86	2.89	2.91	2.92	2.97



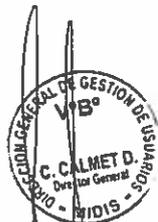
Metodología para la Determinación de la Clasificación Socioeconómica

		Decil del gasto per cápita en el hogar									
		1	2	3	4	5	6	7	8	9	10
Costa Sur Rural	Hogar con jefe con lengua indígena	0.63	0.60	0.54	0.46	0.40	0.37	0.30	0.24	0.16	0.09
	Ratio de hacinamiento	2.26	1.86	1.64	1.45	1.31	1.20	1.03	0.95	0.72	0.60
	Presencia de piso de tierra	0.51	0.40	0.43	0.36	0.24	0.26	0.29	0.22	0.22	0.14
	Usa combustible sólido para cocinar	0.82	0.63	0.57	0.53	0.44	0.39	0.32	0.34	0.26	0.12
	Presencia de techo distinto a concreto	0.83	0.78	0.77	0.79	0.73	0.75	0.68	0.65	0.66	0.54
	Conexión a servicio de agua potable por red pública dentro de la vivienda	0.20	0.26	0.36	0.32	0.37	0.36	0.42	0.33	0.37	0.45
	Conexión a servicio de energía eléctrica por red pública	0.61	0.65	0.78	0.76	0.78	0.73	0.78	0.78	0.85	0.86
Sierra Norte Urbana	Ratio de hacinamiento	1.76	2.49	2.04	1.95	1.74	1.71	1.65	1.43	1.21	0.84
	Presencia de piso de tierra	0.79	0.74	0.72	0.70	0.54	0.42	0.38	0.24	0.17	0.04
	Usa combustible sólido para cocinar	0.89	0.82	0.87	0.71	0.62	0.45	0.46	0.27	0.18	0.04
	Presencia de techo distinto a concreto	1.00	0.93	0.94	0.91	0.86	0.86	0.86	0.77	0.70	0.53
	Máximo año de educación en el hogar	7.24	9.40	9.85	10.06	10.33	10.72	11.00	12.46	13.24	14.34
Sierra Norte Rural	Ratio de hacinamiento	2.46	2.20	2.05	2.01	1.85	1.70	1.44	1.26	1.03	0.88
	Presencia de piso de tierra	0.98	0.95	0.93	0.91	0.88	0.87	0.83	0.73	0.60	0.50
	Usa combustible sólido para cocinar	0.99	0.99	0.98	0.97	0.96	0.94	0.91	0.85	0.68	0.57
	Presencia de techo distinto a concreto	1.00	1.00	0.99	1.00	0.99	0.99	0.99	0.95	0.96	0.96
	El hogar no está conectado a la red pública de desagüe dentro de la vivienda	0.37	0.43	0.49	0.52	0.53	0.58	0.55	0.59	0.67	0.65
	Jefe del Hogar tiene secundaria incompleta	0.96	0.96	0.93	0.91	0.92	0.89	0.86	0.80	0.68	0.51
	Número de bienes per cápita	0.12	0.21	0.31	0.38	0.47	0.55	0.66	0.95	1.36	2.39
	Máximo año de educación en el hogar	6.18	6.72	7.14	7.46	7.56	7.96	7.98	8.40	9.45	10.28
	Número de bienes durables	0.17	0.29	0.40	0.48	0.59	0.69	0.77	1.06	1.44	2.05
	Número de bienes electrodomésticos	0.06	0.12	0.20	0.27	0.34	0.40	0.53	0.75	1.09	1.42
Sierra Centro Urbana	Ratio de hacinamiento	2.48	2.36	2.22	2.05	1.92	1.72	1.57	1.41	1.18	0.81
	Presencia de piso de tierra	0.71	0.53	0.45	0.42	0.35	0.27	0.23	0.18	0.11	0.04
	Usa combustible sólido para cocinar	0.77	0.65	0.54	0.49	0.40	0.35	0.30	0.22	0.14	0.04
	Presencia de techo distinto a concreto	0.95	0.91	0.86	0.83	0.79	0.75	0.70	0.65	0.57	0.44
	Hogar sin ningún servicio público conectado	0.67	0.47	0.36	0.34	0.27	0.22	0.15	0.14	0.09	0.05



Metodología para la Determinación de la Clasificación Socioeconómica

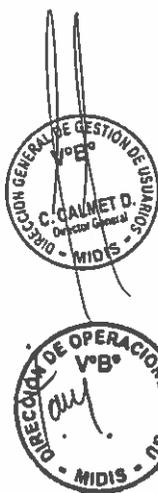
		Decil del gasto per cápita en el hogar									
		1	2	3	4	5	6	7	8	9	10
	JH de hogar con lengua indígena	0.57	0.49	0.47	0.48	0.46	0.43	0.41	0.39	0.33	0.23
	Máximo año de educación en el hogar	7.91	9.43	10.15	10.52	11.04	11.53	12.14	12.68	13.51	14.54
	Máximo año de educación en el hogar	6.43	7.12	7.54	7.46	7.86	8.07	8.40	8.73	9.86	10.90
	Años de educación del jefe del hogar	3.88	4.23	4.55	4.59	4.83	5.16	5.60	6.13	7.71	9.91
	Ratio de hacinamiento	2.02	1.81	1.66	1.53	1.40	1.23	1.14	1.05	0.89	0.70
	Presencia de piso de tierra	0.91	0.89	0.87	0.84	0.83	0.78	0.75	0.66	0.56	0.41
Sierra Centro Rural	Usa combustible sólido para cocinar	0.97	0.96	0.95	0.93	0.93	0.89	0.85	0.78	0.67	0.42
	Número de bienes en el hogar	1.07	1.40	1.69	1.91	2.09	2.44	2.79	3.33	3.98	4.97
	Número de bienes durables en el hogar	0.42	0.54	0.66	0.76	0.83	0.98	1.10	1.39	1.66	2.12
	Número de electrodomésticos	0.14	0.20	0.27	0.33	0.37	0.49	0.60	0.81	1.03	1.40
	Número de bienes per cápita	0.26	0.35	0.43	0.51	0.61	0.73	0.91	1.17	1.74	3.04
	Ratio de hacinamiento	2.15	2.23	1.94	1.91	1.70	1.67	1.54	1.35	1.10	0.76
	Presencia de piso de tierra	0.53	0.36	0.30	0.27	0.26	0.23	0.22	0.16	0.11	0.04
	Usa combustible sólido para cocinar	0.61	0.39	0.36	0.27	0.25	0.21	0.20	0.16	0.11	0.03
Sierra Sur Urbana	Presencia de techo distinto a concreto	0.77	0.63	0.57	0.53	0.52	0.54	0.53	0.51	0.45	0.34
	Hogar sin ningún servicio público conectado	0.32	0.19	0.17	0.14	0.09	0.11	0.10	0.08	0.05	0.03
	JH de hogar con lengua indígena	0.65	0.64	0.62	0.58	0.60	0.60	0.55	0.53	0.48	0.31
	Máximo año de educación en el hogar	8.11	9.65	10.46	11.01	11.33	11.97	12.29	12.81	13.67	14.62
	Máximo año de educación en el hogar	6.46	7.19	7.51	7.88	7.94	8.14	8.63	8.64	9.64	10.81
	Años de educación del jefe del hogar	4.10	4.53	4.92	5.09	5.16	5.47	5.99	6.24	7.48	9.45
	Ratio de hacinamiento	1.97	1.77	1.63	1.52	1.40	1.33	1.21	1.04	0.93	0.67
	Usa combustible sólido para cocinar	0.97	0.95	0.93	0.92	0.91	0.90	0.85	0.82	0.71	0.57
	Número de bienes en el hogar	1.05	1.57	1.79	2.01	2.19	2.33	2.73	3.12	4.01	4.62
	Número de bienes durables en el hogar	0.45	0.66	0.74	0.82	0.88	0.93	1.12	1.27	1.68	1.98
	Número de electrodomésticos	0.13	0.24	0.29	0.38	0.40	0.47	0.59	0.73	1.06	1.29
	Número de bienes per cápita	0.27	0.41	0.50	0.56	0.67	0.75	0.93	1.22	1.67	2.70
	El hogar no está conectado a la red pública de desagüe dentro de la vivienda	0.09	0.11	0.11	0.14	0.13	0.15	0.15	0.18	0.26	0.33
Sierra Sur Rural	Presencia de piso de tierra	0.90	0.85	0.84	0.82	0.81	0.78	0.74	0.71	0.63	0.49



Metodología para la Determinación de la Clasificación Socioeconómica

		Decil del gasto per cápita en el hogar									
		1	2	3	4	5	6	7	8	9	10
Selva Urbana	Ratio de hacinamiento	2.84	2.71	2.39	2.17	1.95	1.73	1.54	1.35	1.14	0.83
	Presencia de piso de tierra	0.56	0.51	0.47	0.41	0.35	0.30	0.24	0.16	0.09	0.04
	Usa combustible sólido para cocinar	0.81	0.70	0.61	0.56	0.49	0.38	0.30	0.20	0.13	0.05
	Presencia de techo distinto a concreto	0.99	0.97	0.96	0.95	0.94	0.91	0.89	0.86	0.79	0.70
	Hogar sin ningún servicio público conectado	0.51	0.44	0.42	0.40	0.32	0.29	0.25	0.19	0.12	0.09
	JH de hogar con lengua indígena	0.16	0.15	0.13	0.13	0.13	0.11	0.09	0.09	0.07	0.05
	Máximo año de educación en el hogar	8.57	9.31	9.88	10.33	10.85	11.15	11.75	12.19	12.91	13.83
Selva Urbana	Ratio de hacinamiento	3.62	2.94	2.57	2.26	1.98	1.71	1.50	1.27	1.04	0.81
	Presencia de piso de tierra	0.48	0.48	0.51	0.48	0.48	0.47	0.44	0.39	0.29	0.25
	Usa combustible sólido para cocinar	0.97	0.95	0.92	0.90	0.87	0.84	0.79	0.73	0.61	0.54
	El hogar está conectado a la red pública de electricidad dentro de la vivienda	0.25	0.31	0.36	0.42	0.48	0.52	0.51	0.61	0.66	0.71
	Número de bienes en el hogar	0.84	1.24	1.66	2.02	2.53	2.83	3.10	3.74	4.55	5.01
	Número de bienes durables en el hogar	0.33	0.49	0.64	0.80	1.01	1.15	1.25	1.55	1.92	2.14
	Número de electrodomésticos	0.14	0.23	0.35	0.45	0.64	0.71	0.82	1.07	1.29	1.51
Lima Metropolitana	Número de bienes per cápita	0.17	0.26	0.38	0.49	0.65	0.79	0.98	1.30	1.87	2.67
	Ratio de hacinamiento	2.29	1.91	1.68	1.53	1.38	1.24	1.11	0.98	0.84	0.59
	Presencia de piso de tierra	0.26	0.17	0.13	0.09	0.07	0.06	0.04	0.03	0.02	0.01
	Usa combustible sólido para cocinar	0.14	0.07	0.05	0.03	0.02	0.02	0.01	0.01	0.00	0.00
	Presencia de techo distinto a concreto	0.70	0.59	0.50	0.43	0.36	0.33	0.26	0.22	0.15	0.08
	Número de bienes per cápita	1.08	1.43	1.71	1.99	2.18	2.47	2.82	3.25	3.92	5.67
	Años de educación del jefe del hogar	8.11	8.63	8.88	9.28	9.61	10.03	10.50	11.06	12.13	13.77

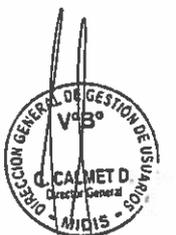
Fuente: ENAHO 2007-2014  
Elaboración: DO-DGGU



Metodología para la Determinación de la Clasificación Socioeconómica

ANEXO A2. Relación de las variables escogidas para segundo factor y su relación con los deciles de gasto per cápita por conglomerado para la base pilada en ENAHO 2007-2014

		Decil del gasto per cápita en el hogar									
		1	2	3	4	5	6	7	8	9	10
Costa Norte Urbana	Número de bienes	3.66	4.48	4.86	5.51	6.01	6.56	7.31	7.81	8.52	9.69
	Número de bienes durables	1.56	1.88	1.96	2.24	2.44	2.64	2.99	3.21	3.58	4.14
	Número de bienes electrodomésticos	0.98	1.32	1.46	1.70	1.89	2.06	2.33	2.46	2.65	3.01
	Número de bienes de telecomunicaciones	0.77	0.90	1.01	1.15	1.27	1.42	1.61	1.80	2.04	2.44
	Número de bienes per cápita	0.74	0.95	1.06	1.27	1.46	1.73	2.02	2.37	2.85	4.15
	Tenencia de licuadora en el hogar	0.38	0.49	0.54	0.61	0.65	0.71	0.77	0.79	0.81	0.86
	Tenencia de plancha eléctrica en el hogar	0.37	0.53	0.58	0.67	0.72	0.77	0.84	0.86	0.86	0.92
Costa Norte Rural	Tenencia de licuadora en el hogar	0.22	0.31	0.42	0.43	0.46	0.57	0.56	0.57	0.55	0.56
	Tenencia de refrigeradora/congeladora en el hogar	0.11	0.15	0.22	0.26	0.30	0.39	0.39	0.42	0.44	0.53
	Tenencia de cocina a gas en el hogar	0.30	0.36	0.46	0.51	0.61	0.66	0.70	0.69	0.74	0.73
	Número de bienes	2.20	2.71	3.43	3.72	4.16	4.76	4.63	4.94	5.29	5.84
	Número de bienes durables	0.82	1.00	1.29	1.43	1.65	1.86	1.88	1.99	2.25	2.49
	Número de bienes electrodomésticos	0.55	0.73	1.03	1.10	1.21	1.53	1.52	1.58	1.69	1.86
	Número de bienes per cápita	0.45	0.59	0.83	1.00	1.24	1.47	1.64	1.97	2.59	3.49
Costa Centro Urbana	Número de bienes	1.04	1.26	1.42	1.50	1.60	1.65	1.84	1.98	2.15	2.33
	Número de bienes durables	4.97	6.09	6.52	6.92	7.31	7.48	8.09	8.53	9.05	9.59
	Número de bienes electrodomésticos	2.15	2.59	2.74	2.88	3.07	3.17	3.43	3.62	3.85	4.16
	Número de bienes de telecomunicaciones	1.46	1.89	2.02	2.18	2.30	2.37	2.57	2.71	2.84	3.00
Costa Centro Rural	Número de bienes per cápita	1.08	1.39	1.60	1.76	2.00	2.16	2.48	2.88	3.45	4.83
	Número de bienes	3.78	4.91	5.26	5.46	5.63	5.90	5.78	5.98	6.13	6.83
	Número de bienes durables	1.64	2.06	2.19	2.30	2.36	2.51	2.49	2.60	2.76	2.87
	Número de bienes electrodomésticos	1.03	1.45	1.59	1.66	1.75	1.91	1.88	2.01	2.03	2.21
Costa Centro Urbana	Número de bienes per cápita	0.86	1.21	1.42	1.63	1.74	1.98	2.07	2.62	3.30	4.73
	Número de bienes	0.79	0.95	1.11	1.23	1.38	1.51	1.73	1.84	2.06	2.47
	Número de bienes durables	3.86	4.81	5.58	6.27	6.88	7.39	8.00	8.55	9.18	10.15
	Número de bienes electrodomésticos	1.72	2.06	2.35	2.67	2.92	3.18	3.45	3.69	3.98	4.45
	Número de bienes de telecomunicaciones	1.00	1.41	1.73	1.97	2.20	2.36	2.55	2.68	2.86	3.17
Costa Centro Rural	Número de bienes per cápita	0.90	1.14	1.44	1.68	1.90	2.27	2.58	3.03	3.77	5.48
	Número de bienes	2.39	3.36	4.10	4.29	5.05	4.70	5.27	5.41	5.98	6.89
	Número de bienes durables	1.05	1.47	1.71	1.81	2.17	1.93	2.30	2.39	2.68	2.98
	Número de bienes electrodomésticos	0.47	0.83	1.22	1.25	1.50	1.36	1.62	1.68	1.79	2.15
	Número de bienes per cápita	0.61	0.91	1.10	1.31	1.73	1.82	2.22	2.71	3.34	4.77

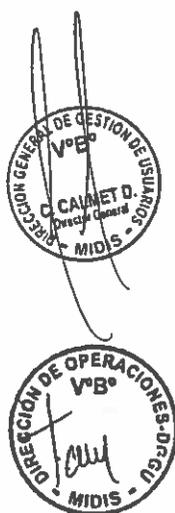


Metodología para la Determinación de la Clasificación Socioeconómica

ANEXO A2. Relación de las variables escogidas para segundo factor y su relación con los deciles de gasto per cápita por conglomerado para la base apilada en ENAHO 2007-2014

		Decil del gasto per cápita en el hogar									
		1	2	3	4	5	6	7	8	9	10
Sierra Norte Urbana	Número de bienes	0.53	0.60	0.75	0.80	0.85	1.02	1.02	1.25	1.48	2.12
	Número de bienes durables	1.63	2.02	2.35	3.03	3.68	4.28	4.94	5.82	6.85	9.05
	Número de bienes electrodomésticos	0.76	0.82	0.86	1.26	1.48	1.67	2.00	2.33	2.80	3.85
	Número de bienes de telecomunicaciones	0.26	0.42	0.59	0.74	1.06	1.23	1.53	1.78	2.10	2.79
	Número de bienes per cápita	0.39	0.44	0.51	0.71	0.87	1.08	1.25	1.59	1.99	3.49
Sierra Centro Urbana	Número de bienes	0.60	0.71	0.86	0.89	0.97	1.08	1.20	1.33	1.58	1.99
	Número de bienes durables	2.33	2.99	3.65	3.97	4.57	5.17	5.66	6.22	7.26	8.41
	Número de bienes electrodomésticos	0.95	1.25	1.46	1.63	1.89	2.10	2.31	2.54	3.00	3.53
	Número de bienes de telecomunicaciones	0.48	0.64	0.87	0.98	1.16	1.41	1.59	1.75	2.12	2.51
	Número de bienes per cápita	0.51	0.68	0.84	0.99	1.14	1.32	1.51	1.80	2.28	3.74
Sierra Sur Urbana	Número de bienes	0.54	0.75	0.91	0.99	1.09	1.21	1.30	1.40	1.70	2.11
	Número de bienes durables	2.50	3.62	4.37	4.94	5.27	5.85	6.17	6.72	7.70	8.97
	Número de bienes electrodomésticos	1.12	1.56	1.88	2.06	2.22	2.45	2.59	2.85	3.28	3.86
	Número de bienes de telecomunicaciones	0.58	0.98	1.28	1.49	1.54	1.74	1.84	2.01	2.31	2.74
	Número de bienes per cápita	0.63	0.89	1.10	1.26	1.42	1.63	1.73	2.06	2.56	4.19
Selva Urbana	Número de bienes	0.57	0.75	0.91	0.99	1.14	1.22	1.39	1.56	1.79	2.18
	Número de bienes durables	2.40	3.35	4.10	4.57	5.15	5.56	6.25	6.93	7.72	8.79
	Número de bienes electrodomésticos	1.01	1.38	1.67	1.84	2.06	2.25	2.54	2.88	3.26	3.75
	Número de bienes de telecomunicaciones	0.61	0.95	1.18	1.38	1.57	1.72	1.96	2.18	2.38	2.69
	Número de bienes per cápita	0.48	0.67	0.85	1.00	1.18	1.38	1.66	2.01	2.52	3.95
Lima Metropolitana	Número de bienes	1.13	1.47	1.67	1.86	1.97	2.14	2.31	2.46	2.74	3.08
	Número de bienes durables	5.41	6.71	7.50	8.12	8.48	8.93	9.40	9.86	10.62	11.49
	Número de bienes electrodomésticos	2.37	2.84	3.16	3.42	3.58	3.83	4.04	4.25	4.61	5.01
	Número de bienes de telecomunicaciones	1.65	2.14	2.43	2.61	2.73	2.84	2.95	3.08	3.26	3.51
	Número de bienes per cápita	1.08	1.43	1.71	1.99	2.18	2.47	2.82	3.25	3.92	5.67

Fuente: ENAHO 2007-2014  
Elaboración: DO-DGGU



ANEXO B. Glosario de Términos

**Clasificación Socioeconómica (CSE) \***: Es una medida de bienestar de las personas y/o sus hogares que comparten condiciones económicas y sociales de manera similar, siendo el resultado de la evaluación mediante la aplicación de la Ficha Socioeconómica Única y/o uso de BDA u otra forma de evaluación aprobada por el MIDIS. La CSE tiene una vigencia de cinco años y se mantendrá vigente hasta su recertificación o actualización.

**Consumo de electricidad per cápita**: Es el resultado del reporte del consumo de electricidad de un hogar dividido por el número de sus miembros. Esta información se obtiene mediante el uso de las bases de datos administrativas o mediante la presentación de los recibos de pago del hogar, validada por la ULE.

En ambos casos el número de registros deberá estar consignado en la FSU. Para los casos de viviendas con un solo hogar se usarán los valores de este medidor. Para el caso de hogares que comparten el consumo, la ULE deberá certificar el número de hogares que obtienen el servicio y dividir el consumo equitativamente.

Además deberá verificar y asignar a los demás hogares que comparten el consumo y que se encuentran en el PGH, así como generar una alerta en el caso de que la información de uno de los hogares que comparten el consumo de energía se incorpore al PGH posteriormente. Cabe enfatizar que esta información puede cambiar la CSE. Asimismo, el análisis e incorporación de la referida información a la evaluación debe corresponder a un período consecutivo de tres meses y correspondiente a los últimos 6 meses (sin contar el vigente).

**Correctamente clasificados**: Es la proporción del total de hogares cuya condición de pobreza ha sido identificada como tal (pobre o no pobre) por un instrumento propuesto en base a la revisión de la literatura económica relevante y las experiencias internacionales en temas de focalización.

**Elegibilidad**: Característica de la población que cumple con los criterios categóricos establecidos por un determinado programa social, el cual puede o no incluir la condición socioeconómica, según Ley de Presupuesto del Sector Público para el Año Fiscal 2013.

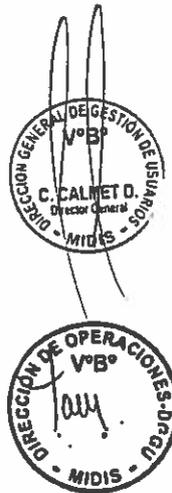
**Error de exclusión**: Es la proporción de hogares en condición de pobreza, según la metodología oficial del INEI, que son identificados como no pobres por un instrumento propuesto en base a la revisión de la literatura económica relevante y las experiencias internacionales en temas de focalización. Equivale a I - sensibilidad.

**Error de inclusión**: Es la proporción de hogares en condición de no pobreza, según la metodología oficial del INEI, que son identificados como pobres por un instrumento propuesto en base a la revisión de la literatura económica relevante y las experiencias internacionales en temas de focalización. Equivale a I - sensibilidad.

**Error total de focalización**: Es la proporción de personas cuya condición de pobreza, según la metodología oficial del INEI, no ha sido identificada correctamente por un instrumento propuesto entre toda la población.

**Especificidad**: Es la proporción de hogares en condición de no pobreza, según la metodología oficial del INEI, que son identificados como tales por un instrumento propuesto en base a la revisión de la literatura económica relevante y las experiencias internacionales en temas de focalización.

**Focalización\***: Conjunto de reglas o instrumentos que permiten identificar a personas o grupos poblacionales en situación de pobreza, vulnerabilidad o exclusión, como potenciales beneficiarios de intervenciones a ser provistas por los programas sociales y subsidios del Estado.



## Metodología para la Determinación de la Clasificación Socioeconómica

**Hogar\***: Conjunto de personas, sean o no parientes, que ocupan en su totalidad o en parte una vivienda o local utilizado como habitación en calidad de residente habitual, que comparten las comidas principales y atienden en común otras necesidades vitales básicas financiadas con un presupuesto común. Por excepción, se considera hogar al constituido por una sola persona.

**Ingreso por planilla pública o privada per cápita**: Es la suma de los ingresos bruto mensual promedio de los miembros del hogar dividido entre el número de miembros que conforman dicho hogar. Dicho valor debe ser obtenido a través de registros administrativos oficiales y deberá corresponder a un período no menor a 3 meses continuos percibidos durante los últimos 6 meses (sin contar el vigente).

**Padrón General de Hogares\***: Es una base de datos nominal de alcance nacional de las personas y/o sus hogares, que permite identificar a los potenciales usuarios de los programas sociales y de subsidios del Estado. Contiene información sobre las características socioeconómicas, la identidad, ubicación, entre otros. Constituye una herramienta principal del Sistema de Focalización de Hogares.

**Pobreza monetaria**: Condición del hogar en el cual el gasto per cápita es insuficiente para adquirir una canasta básica de alimentos y no alimentos.

**Pobreza monetaria extrema**: Condición del hogar en el cual el gasto per cápita no alcanza a cubrir una canasta básica de alimentos.

**Prevalencia de pobreza**: Es la proporción de personas en un determinado ámbito o área que se encuentran en situación de pobreza.

**Programas sociales\***: Intervención pública, previamente planificada y articulada de acciones, prestaciones y beneficios temporales, estructurada a través de objetivos, estrategias, instrumentos y metas, que brinda bienes y/o servicios destinados a lograr un propósito específico sobre personas o poblaciones en situación de pobreza, vulnerabilidad o riesgo social, con el fin de atender una necesidad urgente o revertir un problema que les afecte.

**Punto de corte (o umbral de corte)**: Es el valor específico que tomar el indicador a partir del cual se predice que el hogar se encuentra en condición de pobreza

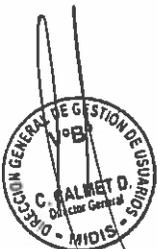
**Recertificación\***: Consiste en la determinación de una nueva certificación de Clasificación Socioeconómica producto de la aplicación de la FSU, el uso de Base de Datos Administrativas u otras formas de evaluación determinadas por el MIDIS. La recertificación efectúa a partir del quinto año desde que una persona obtiene o actualiza su CSE.

**Residente habitual\***: Son aquellas personas que viven permanentemente en el hogar o se encuentran ausentes temporalmente por un período igual o menor a los seis (6) meses.

**Sensibilidad**: Es la proporción de hogares en condición de pobreza, según la metodología oficial del INEI, que son identificados como tales por un instrumento propuesto en base a la revisión de la literatura económica relevante y las experiencias internacionales en temas de focalización.

**Valor predictivo positivo**: Es la proporción de hogares en condición de pobreza, según la metodología oficial del INEI, entre todos los hogares identificados como pobres por un instrumento propuesto en base a la revisión de la literatura económica relevante y las experiencias internacionales en temas de focalización.

**Valor predictivo negativo**: Es la proporción de hogares en condición de no pobreza, según la metodología oficial del INEI, entre todos los hogares identificados como no pobres por un



*Metodología para la Determinación de la Clasificación Socioeconómica*

instrumento propuesto en base a la revisión de la literatura económica relevante y las experiencias internacionales en temas de focalización.

\* Definiciones tomadas de la Directiva N° 1-2015-MIDIS.

