



Resolución Ministerial

Lima, 21 JUL. 2016

N° 151-2016-MIDIS

VISTOS:

El Informe N° 172-2016-MIDIS/VMPES/DGGU de la Dirección General de Gestión de Usuarios; y el Informe N° 054-2016-MIDIS-VMPES/DGGU-DO de la Dirección de Operaciones, hoy Dirección de Investigación y Métodos; 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 Ministerio de Desarrollo e Inclusión Social 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° 107-2015-MIDIS, se aprobó la Metodología para la Determinación de la Clasificación Socioeconómica, la cual viene siendo utilizada por la Dirección de Operaciones de Focalización;

Que, el artículo 4 de la citada Resolución Ministerial N° 107-2015-MIDIS, dispuso 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;

Que, conforme con el artículo 75 del Reglamento de Organización y Funciones del Ministerio de Desarrollo e Inclusión Social, aprobado por Decreto Supremo N° 005-2016-MIDIS, la Dirección de Investigación y Métodos (antes Dirección de Operaciones) es la unidad orgánica de la Dirección General de Gestión de Usuarios que tiene como función, entre otras, la de formular, proponer y coordinar los lineamientos nacionales de focalización, así como, la implementación de las metodologías y los procedimientos del proceso de focalización de acuerdo a las características de cada programa social y/o subsidio del Estado;

Que, el Instituto Nacional de Estadística e Informática (INEI) ha publicado la información correspondiente a la Encuesta Nacional de Hogares (ENAH) 2015;



Que, en dicho contexto, mediante los documentos de Vistos, la Dirección de Investigación y Métodos (antes Dirección de Operaciones) y la Dirección General de Gestión de Usuarios proponen reemplazar la metodología aprobada mediante Resolución Ministerial N° 107-2015-MIDIS, por una nueva;

Que, las citadas Direcciones indican que la nueva metodología debe 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 Padrón General de Hogares (PGH), corresponde a un contexto económico y social diferente al de las estimaciones realizadas en la presente propuesta;

Que, atendiendo a lo expuesto y considerando las competencias y funciones del Ministerio de Desarrollo e Inclusión Social, establecidas en las normas antes reseñadas, se considera pertinente aprobar la nueva Metodología para la Determinación de la Clasificación Socioeconómica, la cual reemplaza a la Metodología aprobada mediante Resolución Ministerial N° 107-2015-MIDIS;

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° 005-2016-MIDIS;

SE RESUELVE:

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

Artículo 2.- Derogar la Resolución Ministerial N° 107-2015-MIDIS, que aprobó la Metodología para la Determinación de la Clasificación Socioeconómica.

Artículo 3.- Disponer que la Metodología para la Determinación de la Clasificación Socioeconómica aprobada en el artículo 1 de la presente resolución, 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 la publicación de la presente resolución y su anexo en el Portal Institucional del Ministerio de Desarrollo e Inclusión Social (www.midis.gob.pe), en la misma fecha de publicación de la presente resolución en el diario oficial El Peruano.

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**



Julio 2016

1. INTRODUCCIÓN

La identificación y selección de usuarios es un proceso de ejecución y mejora continua debido a que no existe una manera automática y única de clasificar a los hogares según su realidad socioeconómica, y a que esta no es estática sino que cambia continuamente. En ese sentido, es importante recalcar que la dinámica que rodea a los hogares ha sufrido cambios muy marcados en los últimos diez años con los fenómenos asociados al crecimiento económico del País, reflejado en el cambio progresivo en las características socioeconómicas de la población, lo cual debe a su vez reflejarse en la elaboración y/o mantenimiento de las herramientas de focalización individual.

El Sistema de Focalización de Hogares (SISFOH), desde su introducción en el 2004, representa un gran esfuerzo por proveer de información socioeconómica a diversos actores de la política social de manera sistemática, con el propósito de guiar las intervenciones focalizadas del Estado en el marco de la política social. En el 2016, el SISFOH es incorporado, mediante Ley N°30435, al Sistema Nacional de Focalización (SINAFO), manteniendo todas sus funciones. Entre ellas, una de las más importantes es la determinación de la clasificación socioeconómica (CSE) de los hogares ante los diversos programas sociales y subsidios del Estado.

2. OBJETIVO

El objetivo es detallar la metodología para la determinación de la clasificación socioeconómica (CSE), a aplicarse por la Dirección de Operaciones de Focalización a cargo de la Dirección General de Gestión de Usuarios del MIDIS, con el fin de que los hogares solicitantes puedan acceder a las intervenciones públicas focalizadas que requieran dicha clasificación dentro de su evaluación de elegibilidad.

3. ALCANCE

La presente metodología es de cumplimiento obligatorio para todos los integrantes del SINAFO, estableciendo la única metodología vigente para la determinación de la CSE.

4. CONTENIDO

4.1. DEFINICIONES

4.1.1. Clasificación Socioeconómica (CSE).- Es una medida de bienestar del hogar. La CSE tiene una vigencia de tres (3) años y se determina con información del hogar que se encuentre en las Bases de Datos Administrativas (BDA), así como aquella que es recolectada mediante instrumentos aprobados por el MIDIS.

La CSE se registra en el Padrón General de Hogares (PGH) con las siguientes variables:

- | | | |
|-------|---------------------------------|---|
| (i) | Código de hogar | Código que identifica de manera única a cada hogar. |
| (ii) | Código de CSE: | Código que identifica de manera única la CSE del hogar. Debe ser empleado por los agentes responsables de los programas sociales y subsidios del Estado al momento de afiliar a sus usuarios. |
| (iii) | Nivel de la CSE | Nivel socioeconómico asignado al hogar, pudiendo ser: (i) pobre extremo, (ii) pobre y (iii) no pobre. |
| (iv) | Fecha de inicio de la vigencia: | Indica la fecha en la cual la CSE asignada al hogar comienza a tener validez en el PGH. |
| (v) | Fecha de fin de la vigencia: | Indica la fecha en la cual la CSE asignada al hogar deja de tener validez en el PGH. |

- | | | |
|--------|---|--|
| (vi) | Estado de la vigencia: | Indica si la CSE se encuentra: (i) vigente, (ii) próxima a caducar; (iii) caduca, o (iv) cancelada |
| (vii) | Norma que aprueba el Método de cálculo de la CSE: | Indica la norma con la que se aprueba la metodología para determinar la CSE. |
| (viii) | Norma que establece la Vigencia de la CSE: | Indica la norma con la que se aprueba el plazo de vigencia de la CSE. |

4.1.2. Determinación de la CSE.- Es el proceso mediante el cual, aplicando la metodología vigente aprobada por el MIDIS, se asigna la CSE a un hogar.

4.1.3. Empadronamiento de hogares.- Procedimiento que consiste en registrar información del hogar, de sus integrantes y de su dirección domiciliaria, en el PGH, de acuerdo con las disposiciones aprobadas por el MIDIS.

4.1.4. 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. Comparten las comidas principales y atienden entre sus integrantes otras necesidades vitales básicas financiadas con un presupuesto común.

Esta definición incluye a los integrantes que, sin ser residentes habituales, dependen del presupuesto del hogar, siempre y cuando cursen estudios y tengan hasta 22 años de edad. Excepcionalmente, se podrá considerar como hogar sin vivienda a los conformados por una o más personas en situación de calle o albergues públicos.

4.1.5. Instrumentos de recojo de datos.- Son herramientas para recolectar datos directamente de la vivienda, el hogar y las personas residentes habituales, así como otras características, con el objeto de determinar y verificar la CSE. Entre ellos se encuentran el Formato de Solicitud (S100) y la Ficha Socioeconómica Única (FSU). La suscripción de estos instrumentos por parte del solicitante o informante les da el carácter de una declaración jurada.

4.1.6. Padrón General de Hogares (PGH).- Es una base datos de hogares y sus integrantes, el cual incluye la composición del hogar, la identidad de los integrantes y su clasificación socioeconómica.

4.1.7. Residente habitual.- Es aquella persona que vive permanentemente en el domicilio del hogar, incluyendo a aquellas que se encuentran ausentes temporalmente por un periodo igual o menor a los seis (6) meses.

4.1.8. Solicitud de CSE.- Es el pedido formal que realiza una persona, suscribiendo el o los formatos de solicitud vigentes, a fin que se evalúe el estado de la CSE de su hogar en el PGH.

4.2. CONTENIDOS GENERALES

4.2.1. Pobreza y pobreza monetaria

Se considera pobreza a la condición de falta de bienestar producida por la incapacidad de satisfacer las necesidades básicas. En el Perú, como en la mayoría de países latinoamericanos, la pobreza se mide desde un enfoque monetario, según el cual el bienestar es indicado por el gasto (el consumo real)¹.

De esta forma, se consideran pobres monetarios a las personas que residen en hogares cuyo gasto per cápita es inferior a la línea de pobreza, la cual es el gasto mínimo para adquirir una canasta básica de consumo, compuesta por la canasta básica de alimentos y no alimentos (un conjunto de bienes y servicios considerados necesarios para tener

¹ Dicha elaboración permite diseñar de políticas públicas orientadas directamente a la reducción de la incidencia de la pobreza.

condiciones mínimas de vida tales como vivienda, vestido, educación, salud, transporte, etc.). Adicionalmente, son pobres extremos aquellas personas que integran hogares cuyos gastos per cápita están por debajo del costo de la canasta básica de alimentos². La presente metodología de clasificación socioeconómica confronta su ajuste predicho a este enfoque de medición de la pobreza.

4.2.2. 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. 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.

Asimismo, a fin de incorporar las diferencias en las dinámicas territoriales, la evaluación de bases de datos administrativas 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, la gradiente de urbanidad adicional considera 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, respecto del cálculo del IFH, al ser necesaria una mayor precisión en los cálculos, se desagrega el territorio nacional en quince conglomerados que articulan información geográfica y social, sobre la base de patrones de consumo diferenciados. En la siguiente sección se presentan a detalle dicha distinción.

Para evaluar la precisión de las predicciones estimadas por los instrumentos que forma la presente metodología, se hizo uso de las siguientes herramientas:

4.2.3. Método logístico

Es un modelo de regresión estadística en la cual la variable dependiente es binaria. Este modelo es usado para estimar la probabilidad de que se cumpla o no una condición basado en uno o más predictores (variables independientes). Bajo esta lógica, la condición de pobre o no es la variable dicotómica, las variables de registros administrativos o las proyecciones de las variables socioeconómicas calculadas por un índice son potenciales predictores a probar.

4.2.4. Análisis ROC (Receiver Operating Characteristic)

El análisis ROC permite evaluar un modelo estadístico en cuanto a su capacidad para discriminar correctamente entre la población calificada como “tiene la condición” y “no tiene la condición”. Este análisis se aplica sobre las probabilidades predichas por un modelo logístico previamente construido, comparando dichas predicciones contra la verdadera situación de la condición. Para nuestro caso, la condición predicha es “hogar en condición de pobreza” según lo informado por la ENAHO del INEI.

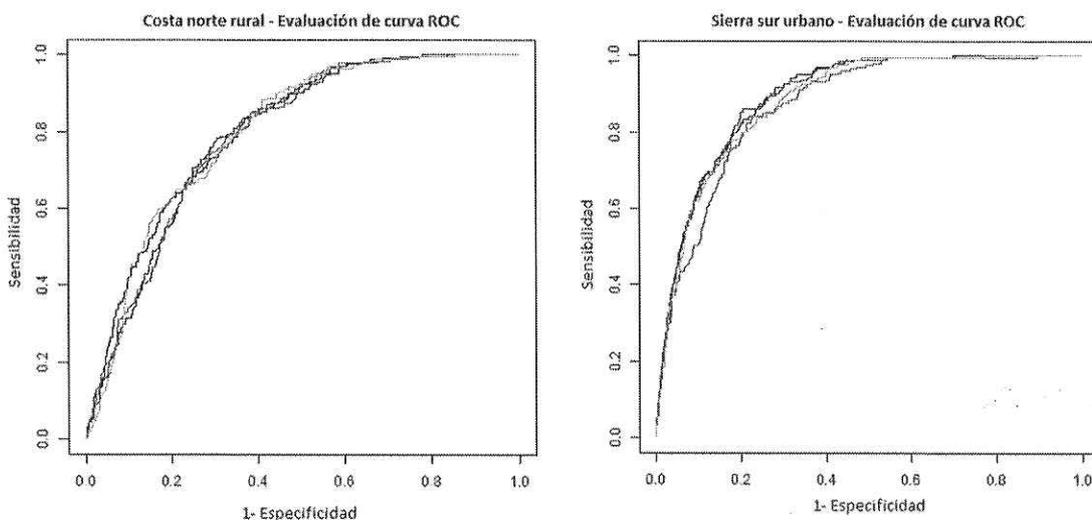
Se denomina “curva ROC” a la representación gráfica de los errores de predicción (sensibilidad, especificidad) construidos de esta manera. La proporción de hogares pobres correctamente predichos del total de hogares pobres se denomina *sensibilidad* (su complemento es el *error de exclusión*), y la proporción de hogares no pobres correctamente predichos como tales del total de hogares no pobres se denomina

² INEI (2015). Informe Técnico. Evolución de la Pobreza Monetaria 2009-2015
https://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitaless/Est/Lib1347/libro.pdf

especificidad (cuyo complemento es el *error de exclusión*). En cualquier caso, la disminución de uno de estos dos errores (usando umbrales más altos o más bajo) conlleva el aumento del otro. Dicho intercambio, no es proporcional para todos los puntos de la gráfica, y es posible encontrar lugares donde se puede ganar precisión en un sentido con mínimo costo en el otro, por lo cual el gráfico resultante no es una recta, sino una curva. El área de esta curva toma usualmente un valor entre 0.5 (recta plana, el peor valor) y 1 (curva cercana al vértice superior derecho, el mejor valor).

Todos los instrumentos cuantitativos fueron evaluados mediante este análisis para buscar posibles puntos de corte, así como los valores del error para dichos puntos de corte, teniendo como variable resultado la condición de pobreza de un hogar según lo informado por la ENAHO.

Gráfico 1: Análisis de Curvas ROC



4.2.5. Marco teórico de Bayes

El teorema de Bayes y la inferencia bayesiana presentan un marco teórico estadístico para calcular la probabilidad de que un elemento de un conjunto presente realmente 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 (*probabilidad pre-test*) 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 (*probabilidad post-test*) cercana a 100% entonces es muy probable de ser pobre, mientras que el 0% indica con mucha certeza que es no pobre.

De esta manera, la metodología propuesta arroja una clasificación socioeconómica que resulta creíble o no, dependiendo de la zona geográfica en la que resida el hogar. A manera de ejemplo, una clasificación de pobreza no resulta creíble para una zona geográfica cuya tasa de pobreza sea muy baja, mientras que, al contrario, una clasificación de no pobreza tampoco es creíble para una zona geográfica cuya tasa de pobreza sea muy alta. Esto permite usar la evidencia real del contexto de la ocurrencia de la pobreza para ponderar la probabilidad de presentar realmente la CSE.

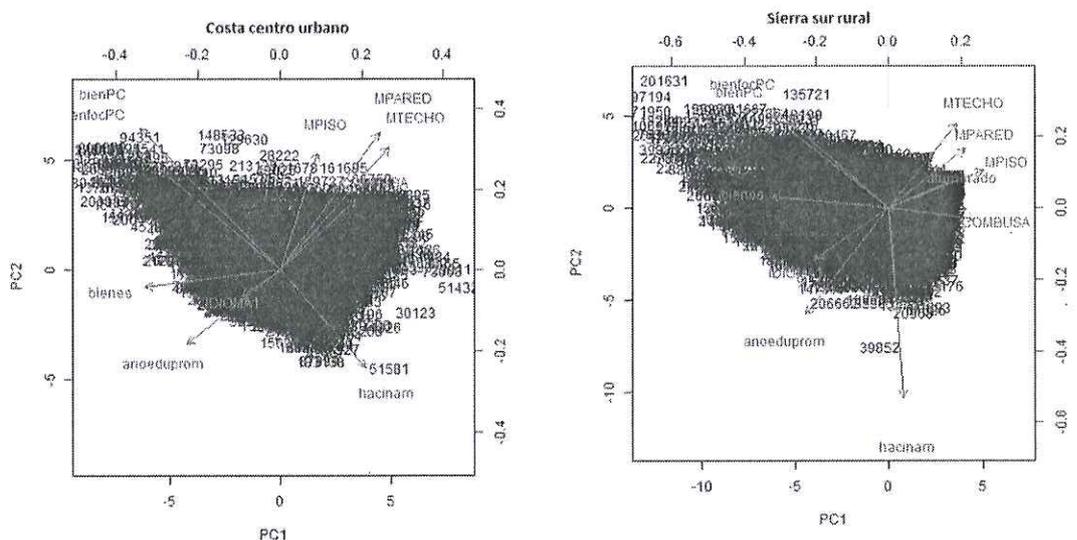


4.2.6. Análisis de Componentes Principales (PCA, Principal Component Analysis)

El PCA es un método estadístico que busca reducir las dimensiones de un conjunto de variables aleatorias, conservando la mayor parte de la información expresada en la variabilidad. Para ello se realiza una transformación ortogonal, con la cual se pasa de un conjunto de variables posiblemente correlacionadas a un conjunto de variables no correlacionadas.

Estas reciben el nombre de componentes principales y se encuentran ordenadas según su importancia en la explicación de la varianza. Los resultados del PCA se interpretan a partir de los *proxy means test* (PMT), los cuales muestran información simultáneamente de las variables involucradas en describir a los hogares. De esta forma, se busca construir un único índice que sea la proyección de la transformación lineal de las variables involucradas en dichos componentes y que se acerque a un indicador más objetivo de la pobreza, como lo es el gasto del hogar. En ese sentido, suele utilizarse el primer componente obtenido para la estimación de los parámetros del Índice de Focalización de Hogares (IFH).

Gráfico 2: Análisis de Componentes Principales



Si bien no hay un método claro para definir qué 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 y (ii) lograr niveles altos de áreas bajo las curvas ROC del primer componente, el criterio más utilizado en la evaluación de Proxy Mean Tests; y (iii) obtener un factor (producto de PCA) que se distribuya de forma normal.

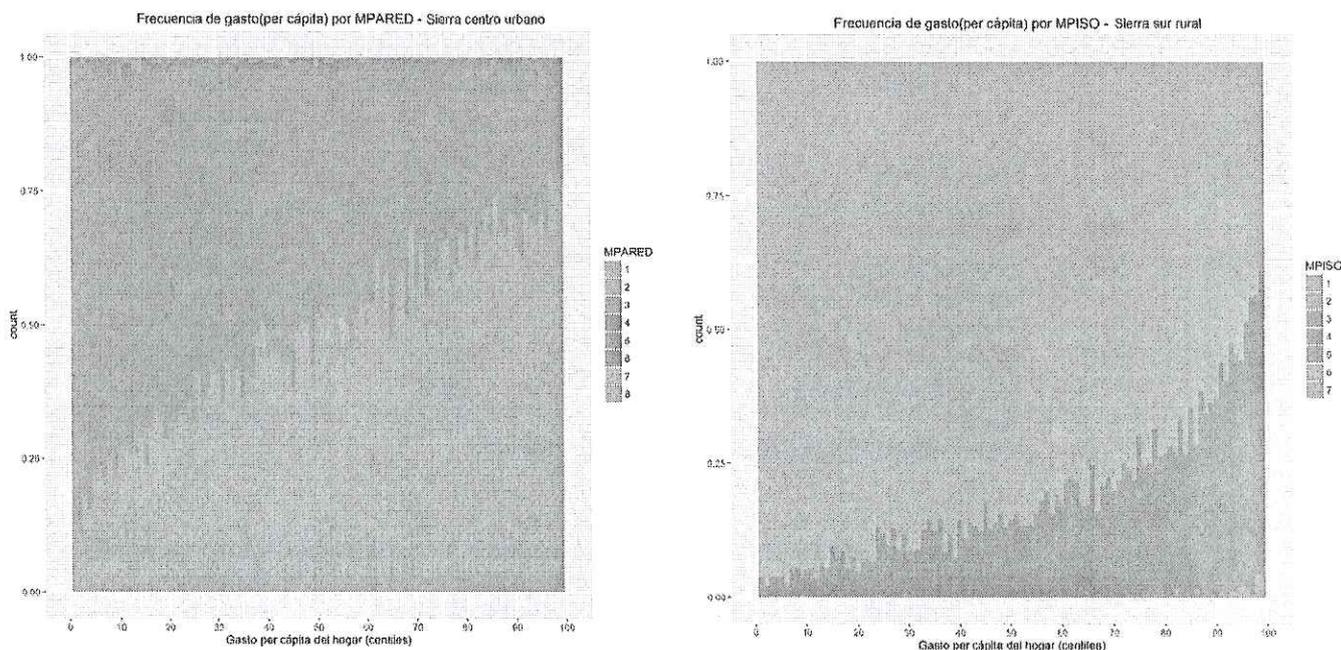
4.2.7. Re-escalamiento de variables nominales

Durante la elaboración de la Metodología para la Determinación de la Clasificación Socioeconómica se evidenció el comportamiento heterogéneo de las variables socioeconómicas en cuanto a su correlación con el consumo del hogar. Así, algunas características pudieron ser representativas de menor gasto en algunos ámbitos y, al mismo tiempo, ser representativa de mayor gasto en otros.

Debido a ello, y al mejor comportamiento de variables ordinales en lugar de aquellas dicotómicas en los modelos estadísticos compuestos, se escalan las variables de acuerdo a su relación con el gasto para cada ámbito, de manera que un mayor valor correspondiese a mayor pobreza. Para la inspección se dividió el gasto per cápita por centiles, para evidenciar el comportamiento de cada uno de los centiles por variable. Por ejemplo, en el gráfico 3 se observa una relación inversa casi perfecta entre el piso de tierra (categoría 6) y un mayor gasto. Asimismo, aquellos pisos construidos con material noble (categorías 1, 2 y 3) presentan una relación positiva directa con el gasto.



Gráfico 3: Cruce de variables por centiles de gasto per cápita del hogar



4.3. CONTENIDOS ESPECÍFICOS

A continuación se desarrollan los pasos que componen la metodología para la determinación de la CSE.

4.3.1. Paso 1: Evaluación de la afiliación a seguros de salud privados

Evaluación de la afiliación a un seguro privado de salud (SPS) o el acceso a EPS (empresa prestadora de salud) por parte del jefe del hogar o el/la cónyuge. Al respecto, al analizar la relación entre la situación de no pobreza y la tenencia de un seguro privado de salud o la afiliación a una entidad prestadora de salud (EPS), se comprueba que la mayoría de la población con esta característica no se encuentra en situación de pobreza, en todos los ámbitos geográficos (ver tabla 1). Por lo tanto, es una herramienta efectiva para identificar a los hogares no pobres.

Los hogares con esta característica serán considerados como no pobres y no siguen al paso (2).

4.3.2. Paso 2: Evaluación de la tenencia de vehículos

El análisis por ámbito de la tenencia de un vehículo tipo auto o camioneta permite evidenciar, a partir de los valores predictivos negativos, que al igual que el caso anterior la tenencia de estos vehículos muestra una alta capacidad de descarte de la pobreza para todos los ámbitos (ver tabla 2). Dicha información se obtiene a través de las bases de datos administrativas de la Superintendencia Nacional de los Registros Públicos (SUNARP).

De esta manera, **los hogares en los cuales alguno de sus miembros tenga registrada la posesión de un automóvil o camioneta serán considerados como no pobres y no siguen al paso 3.**



Tabla 1. Predicción de la variable "Jefe del hogar o cónyuge afiliado a EPS/SPS" versus la condición de pobreza, por ámbito

Ámbito	Sensibilidad	Especificidad	Valor predictivo positivo	Valor predictivo negativo
Costa Norte	100.0%	2.5%	16.6%	100.0%
Costa Centro	99.6%	1.2%	9.8%	96.2%
Costa Sur	100.00%	2.8%	9.6%	100.0%
Sierra Norte	99.9%	1.8%	47.3%	95.8%
Sierra Centro	100.0%	0.6%	27.6%	100.0%
Sierra Sur	100.0%	1.1%	21.0%	100.0%
Selva	100.0%	1.2%	22.7%	100.0%
Lima Metropolitana	99.2%	9.5%	8.7%	99.3%
Perú	99.9%	4.3%	18.5%	99.3%

Elaboración: VMPEP/DGGU-DO. Fuente: ENAHO 2015

Tabla 2. Predicción de la variable "Posesión de auto o camioneta" versus la condición de pobreza, por ámbito

Ámbito	Sensibilidad	Especificidad	Valor predictivo positivo	Valor predictivo negativo
Costa Norte	99.2%	7.2%	17.2%	97.8%
Costa Centro	97.3%	10.1%	9.9%	97.3%
Costa Sur	94.4%	21.2%	10.9%	97.4%
Sierra Norte	99.7%	6.2%	48.4%	96.1%
Sierra Centro	99.1%	8.8%	29.1%	96.4%
Sierra Sur	98.6%	14.0%	23.2%	97.4%
Selva	99.7%	3.9%	23.2%	97.9%
Lima Metropolitana	98.5%	18.2%	9.3%	99.3%
Perú	99.0%	12.3%	19.7%	98.3%

Elaboración: VMPEP/DGGU-DO. Fuente: ENAHO 2015

4.3.3. Paso 3: Evaluación de ingresos del hogar

La evaluación de ingresos se realiza a partir de la información registrada por la SUNAT respecto a las rentas de primera, tercera, cuarta y quinta categoría. Respecto a los tipos de renta³, distinguimos a aquellas de primera categoría, que corresponden al alquiler de un predio, inmueble y/o vehículo por parte del propietario o subarrendatario.

En segundo lugar, las rentas de tercera categoría se refieren a aquellas obtenidas por la realización de actividades empresariales desarrolladas por personas naturales y jurídicas, generalmente producidas por la participación conjunta de la inversión del capital y el trabajo. En tercer lugar, las rentas de cuarta categoría (o de trabajo independiente) son los ingresos que provienen de realizar cualquier trabajo individual, de una profesión, arte, ciencia u oficio cuyo cobro se realiza sin tener relación de dependencia contractual. Estas incluyen también las contraprestaciones obtenidas por quienes prestan servicios bajo la modalidad de Contrato Administrativo de Servicios (CAS). Finalmente, las rentas de quinta categoría (o de trabajo dependiente) se definen como las rentas generadas por los trabajadores que laboran para un empleador.

Para la selección de umbrales en esta evaluación de ingresos del hogar, se tomó en cuenta que el uso temprano de estos instrumentos para el descarte de hogares no pobres exigía altos valores de sensibilidad que mejoren el rendimiento del índice de focalización de hogares (IFH) para las etapas posteriores de la clasificación socioeconómica. Por ello se evaluó el rendimiento predictivo para las frecuencias de

³ SUNAT (2016), obtenido de <http://www.orientacion.sunat.gob.pe/index.php/personas-menu/impuesto-a-la-renta-personas-ultimo>

5%. Se optó por los puntos de corte de valores de consumo de electricidad per cápita de S/ 24, S/ 13 y S/ 9 para los ámbitos de Lima Metropolitana, resto urbano y rural, respectivamente. **Aquellos hogares por encima de estos valores son considerados no pobres y no siguen al paso 5**, debido a que tienen una probabilidad muy baja de ser pobres (ver tabla 4).

Tabla 4. Predicción del nivel de consumo de electricidad per cápita del hogar versus la condición de pobreza, por ámbito

Ámbito	Umbrales (en S/)	Sensibilidad	Especificidad	Valor predictivo positivo	Valor predictivo negativo
Costa Norte Urbano	13.00	98.5%	20.7%	37.2%	96.8%
Costa Norte Rural	9.00	96.9%	31.2%	23.3%	97.9%
Costa Centro Urbano	13.00	98.9%	10.8%	46.7%	92.6%
Costa Centro Rural	9.00	98.8%	9.6%	43.4%	91.7%
Costa Sur Urbano	13.00	99.1%	12.8%	45.2%	95.2%
Costa Sur Rural	9.00	98.5%	11.5%	41.9%	92.0%
Sierra Norte Urbano	13.00	98.7%	20.4%	27.1%	98.2%
Sierra Norte Rural	9.00	87.3%	61.5%	7.0%	99.3%
Sierra Centro Urbano	13.00	99.0%	21.1%	27.0%	98.6%
Sierra Centro Rural	9.00	92.3%	45.8%	11.6%	98.7%
Sierra Sur Urbano	13.00	99.5%	14.6%	34.6%	98.4%
Sierra Sur Rural	9.00	93.1%	39.2%	11.8%	98.5%
Selva Urbana	13.00	99.0%	18.7%	29.9%	98.2%
Selva Rural	9.00	97.9%	35.9%	10.9%	99.5%
Lima Metropolitana	24.00	99.2%	11.9%	30.4%	97.5%
Perú		98.4%	27.8%	28.5%	98.4%

Elaboración: VMPES/DGGU-DO. Fuente: ENAHO 2015

4.3.5. Paso 5: Cálculo del Índice de Focalización de Hogares

Los conglomerados considerados para el cálculo del IFH 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 la tabla 5. Cabe precisar que anualmente la ENAHO recolecta información de al menos 20,000 hogares.

Tabla 5. Muestra Apilada por Conglomerado Geográfico ENAHO 2007-2015

Dominio Geográfico	U/R	Muestra
Costa Norte	Rural	6,394
	Urbana	24,731
Costa Centro	Rural	3,502
	Urbana	14,902
Costa Sur	Rural	2,089
	Urbana	10,384
Sierra Norte	Rural	11,092
	Urbana	3,327
Sierra Centro	Rural	26,154
	Urbana	17,477
Sierra Sur	Rural	18,661
	Urbana	15,361
Selva	Rural	21,851
	Urbana	26,066
Lima Metropolitana		28,275
TOTAL		230,266

Elaboración: VMPES/DGGU-DO. Fuente: ENAHO 2007-2015

pobreza que se esperaría encontrar en el campo, a condición de que la predicción negativa/el valor predictivo negativo se mantenga sobre el 90%.

Una segunda consideración importante es el hecho de que la información de planillas públicas y privadas presenta solo los ingresos de trabajadores del sector formal de la economía. Esta, según cifras del 2012, comprende sólo al 26% de la Población Económicamente Activa (PEA) a nivel nacional: 40%, 33% y 5% de Lima Metropolitana, el resto urbano y el ámbito rural, respectivamente.

El cálculo de umbrales para este modelo se dificulta por estos bajos porcentajes de población en el sector formal, dado que estos podrían contener grupos de hogares pobres y no pobres bastante semejantes. Por ello, por ejemplo, familias pobres podría pasar por pobres solo por el hecho de pertenecer al sector formal. Para evitar esta distorsión, sobre todo en los ámbitos resto urbano y rural, se emplean valores de corte mayores a la línea de pobreza, puesto que esta presenta una sensibilidad baja.

Según lo expuesto, se comparan los ingresos mensuales promedio per cápita del hogar contra los puntos de corte de S/ 631 para Lima Metropolitana, S/ 543 para el resto urbano y S/ 358 para el ámbito rural. **Aquellos hogares por encima de estos valores son considerados no pobres y no siguen al paso 4**, debido a que tienen una probabilidad muy baja de ser pobres (ver tabla 3).

Tabla 3. Predicción del nivel de ingreso per cápita del hogar versus la condición de pobreza, por ámbito

Ámbito	Umbrales (en S/)	Sensibilidad	Especificidad	Valor predictivo positivo	Valor predictivo negativo
Costa Norte Urbano	543	99.4%	19.7%	31.5%	98.9%
Costa Norte Rural	358	96.8%	32.1%	26.7%	97.5%
Costa Centro Urbano	543	99.7%	9.6%	35.1%	98.3%
Costa Centro Rural	358	98.4%	9.9%	46.8%	88.3%
Costa Sur Urbano	543	99.3%	13.1%	45.6%	96.4%
Costa Sur Rural	358	97.3%	12.8%	54.4%	81.3%
Sierra Norte Urbano	543	98.6%	23.3%	39.2%	97.0%
Sierra Norte Rural	358	88.6%	63.9%	16.6%	98.6%
Sierra Centro Urbano	543	99.4%	23.5%	36.1%	99.0%
Sierra Centro Rural	358	95.6%	47.5%	17.1%	99.0%
Sierra Sur Urbano	543	99.6%	16.1%	41.9%	98.4%
Sierra Sur Rural	358	95.6%	41.2%	18.8%	98.5%
Selva Urbana	543	99.3%	21.2%	40.0%	98.4%
Selva Rural	358	96.9%	40.1%	26.3%	98.3%
Lima Metropolitana	631	99.0%	13.7%	42.3%	95.5%
Perú		98.6%	29.4%	34.0%	98.3%

Elaboración: VMPE/DGGU-DO. Fuente: ENAHO 2015

4.3.4. Paso 4: Evaluación por el consumo en servicios de electricidad

Para esta evaluación, se emplea como indicador el consumo per-cápita del hogar en el servicio de alumbrado eléctrico de red pública. Para evaluar el consumo en servicios de electricidad la información se obtiene a través de las bases administrativas del Organismo Supervisor de la Inversión en Energía y Minería (OSINERGMIN).

El análisis de este indicador reporta un uso bastante extendido de dicho servicio público. Para el 2015, el 90% de hogares tratados por la ENAHO cuenta con alumbrado eléctrico. Al desagregar la información por ámbitos, se obtiene niveles de acceso de 99.7%, 98.5% y 76.3% para las viviendas en Lima Metropolitana, el resto urbano y el ámbito rural, respectivamente.

Con estas cifras, se evalúan los puntos de corte vigentes, para observar la capacidad de los mismo de predecir, buscando valores de predicción post-test negativos inferiores a



Para la construcción del IFH, se han evaluado todas las variables de la ficha socioeconómica única (FSU), las cuales se han dividido en cuatro categorías: (i) bienes del hogar, (ii) servicios básicos del hogar, (iii) calidad de la vivienda y (iv) población.

En base al análisis gráfico, las categorías que aportan positivamente al IFH corresponden a las variables que están asociadas a mayores niveles de precariedad mientras que aquellas que aportan negativamente se asocian a un menor nivel de precariedad.

Luego de excluir las variables que no detentan una correlación fuerte con la pobreza, se construye el IFH según el análisis por percentiles de las variables restantes para los quince ámbitos. Según el análisis de componentes principales, con las variables seleccionadas, no fue posible determinar un componente que pueda explicar más del 50% de la varianza.

Tabla 6. Área debajo de la Curva ROC por Factor y Conglomerado Geográfico ENAHO 2007-2015

Dominio Geográfico	U/R	ROC
Costa Norte	Urbano	83.9%
	Rural	78.8%
Costa Centro	Urbano	85.1%
	Rural	84.5%
Costa Sur	Urbano	87.1%
	Rural	83.9%
Sierra Norte	Urbano	86.5%
	Rural	71.4%
Sierra Centro	Urbano	86.4%
	Rural	72.6%
Sierra Sur	Urbano	89.3%
	Rural	72.1%
Selva	Urbano	87.2%
	Rural	73.9%
Lima Metropolitana		86.7%

Elaboración: DO-DGGU. Fuente: ENAHO 2007-2015

A pesar de ello, en el tratamiento del modelo fue posible obtener curvas ROC de más del 70%; lo cual refleja una capacidad predictiva satisfactoria de la pobreza. La tabla 6 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 utilizar dichos modelos como instrumento para identificar a la población en situación de pobreza.

Finalmente, la selección de los umbrales y modelos del IFH para cada ámbito geográfico se realizan en simultáneo. Para ello se toman en cuenta las siguientes consideraciones: conseguir un valor ROC superior al 0.75, recoger la mayor cantidad de dimensiones socioeconómicas, tener una probabilidad post-test negativo de pobreza menor al 10% y una sensibilidad mayor a 65%, todo ello procurando la mejor probabilidad post-test positiva, según el marco teórico de Bayes.

Así, los umbrales de corte por conglomerado son los siguientes (ver tabla 7), para los modelos predictivos (ver tabla 8).

Tabla 7. Umbrales de Corte de los Factores por Conglomerado-IFH 2016

Dominio Geográfico	U/R	Pobreza Total	Pobreza Extrema
Costa Norte	Urbano	0.92	3.36
	Rural	-0.34	1.43
Costa Centro	Urbano	0.96	4.47
	Rural	0.95	1.97
Costa Sur	Urbano	1.05	3.66
	Rural	0.92	2.23
Sierra Norte	Urbano	0.95	2.35
	Rural	-1.25	0.54
Sierra Centro	Urbano	1.23	3.02
	Rural	-0.69	1.03
Sierra Sur	Urbano	1.61	3.56
	Rural	-0.18	1.03
Selva	Urbano	1.22	2.65
	Rural	-0.48	1.22
Lima Metropolitana		1.13	2.95

Elaboración: DO-DGGU. Fuente: ENAHO 2007-2015

Tabla 8. Predicción del Índice de Focalización de Hogares (IFH) versus la condición de pobreza, por ámbito

Ámbito	Sensibilidad	Especificidad	Valor predictivo positivo	Valor predictivo negativo
Costa Norte Urbano	74.8%	76.3%	31.7%	95.4%
Costa Norte Rural	84.1%	61.1%	44.1%	91.3%
Costa Centro Urbano	70.2%	80.2%	21%	97.3%
Costa Centro Rural	70.0%	83.8%	19.8%	98%
Costa Sur Urbano	81.5%	78.1%	24.6%	98%
Costa Sur Rural	65.0%	82.5%	22.8%	96.7%
Sierra Norte Urbano	78.4%	81.7%	43.9%	95.4%
Sierra Norte Rural	86.2%	43.6%	66.7%	70.8%
Sierra Centro Urbano	73.6%	81.4%	43.9%	94%
Sierra Centro Rural	82.2%	53.8%	51%	83.9%
Sierra Sur Urbano	77.9%	83.1%	33.1%	97.2%
Sierra Sur Rural	76.5%	60.3%	49%	83.7%
Selva Urbana	69.3%	84.7%	41.3%	94.7%
Selva Rural	81.4%	52.6%	44.6%	85.8%
Lima Metropolitana	69.6%	84.1%	24.7%	97.4%

Elaboración: VMPES/DGGU-DO. Fuente: ENAHO 2015



Anexo 1

Coeficientes de las variables por conglomerado del área urbana- IFH 2016								
	Costa			Sierra			Selva	Lima
	Norte	Centro	Sur	Norte	Centro	Sur		
Vivienda								
1. Ratio de hacinamiento	0.18	0.24	0.27	0.21	0.17	0.20	0.25	0.29
2. Material del Piso	0.32	0.37	0.36	0.35	0.34	0.35	0.33	0.38
3. Material del Techo	0.24	0.33	0.28	0.22	0.25	0.31	0.23	0.32
4. Material de las Paredes	0.26	0.30	0.23	0.27	0.27	0.30		
Servicios								
1. Combustible usado para cocinar		0.21		0.33	0.29	0.28	0.33	0.16
2. Abastecimiento de agua	0.19			0.19				
3. Conexión de los servicios higiénicos	0.25			0.24	0.24		0.27	
4. Tipo de alumbrado				0.04	0.14			
Bienes								
1. bienes per-cápita	-0.34	-0.39	-0.39	-0.41	-0.35	-0.35	-0.41	-0.41
2. bienes focalizados per-cápita	-0.35	-0.42	-0.43	-0.42	-0.37	-0.37	-0.41	-0.45
3. bienes en el hogar	-0.40	-0.39	-0.41	-0.41	-0.38	-0.37	-0.40	-0.40
4. Licuadora	-0.27							
5. Refrigeradora/congeladora	-0.30							
Población								
1. Años de educación de mayores de edad	-0.28	-0.27	-0.31		-0.33	-0.30	-0.31	-0.33
2. Idioma del jefe del hogar		-0.11	-0.25		-0.22	-0.23		
3. Ocupación del jefe del hogar						0.15		
4. Estado Civil del jefe del hogar						0.11		

Fuente: ENAHO 2007-2015. Elaboración: DO-DGGU



Coeficientes de las variables por conglomerado del área rural- IFH 2016

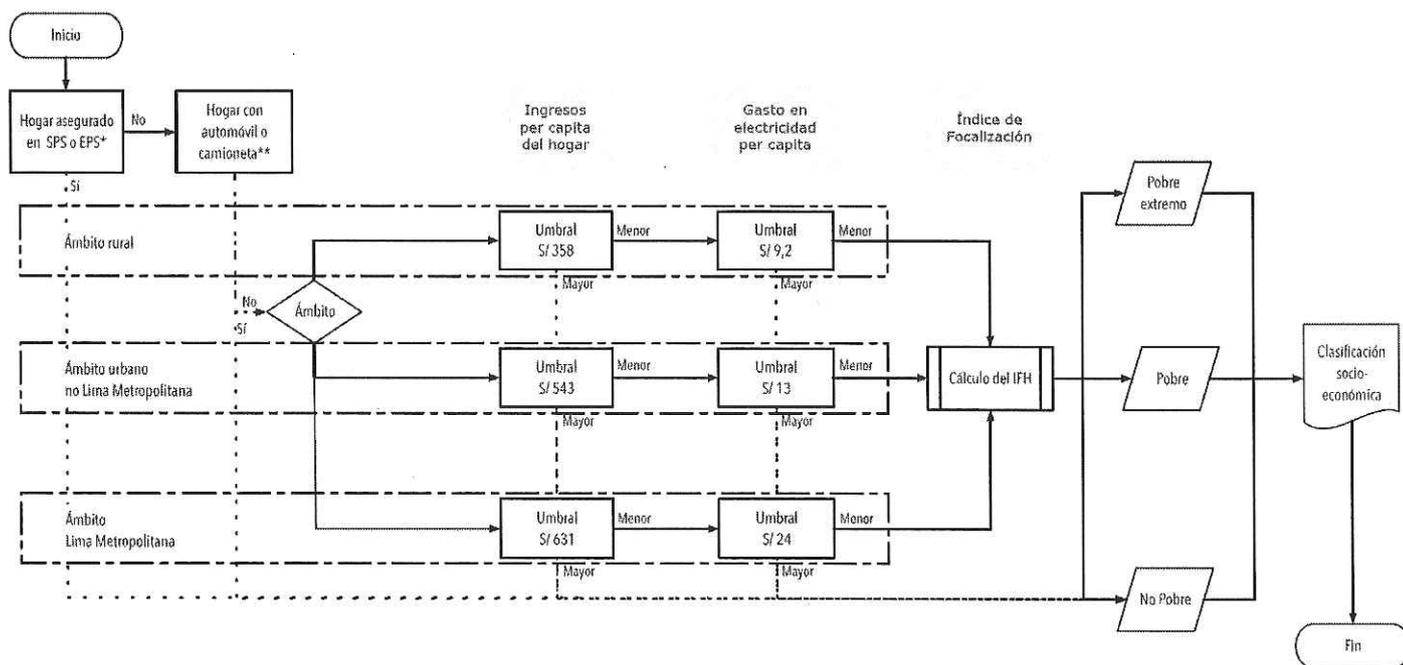
	Costa			Sierra			Selva
	Norte	Centro	Sur	Norte	Centro	Sur	
Vivienda							
1. Ratio de hacinamiento	0.18	0.27	0.23	0.18	0.08	0.06	0.21
2. Material del Piso	0.30	0.40	0.34	0.33	0.30	0.30	0.31
3. Material del Techo	0.10	0.24				0.20	
4. Material de las Paredes	0.03				0.17	0.24	0.20
Servicios							
1. Combustible usado para cocinar	0.33	0.35	0.34	0.33	0.27	0.28	0.32
2. Abastecimiento de agua	0.21				0.19	0.15	0.23
3. Conexión de los servicios higiénicos	0.22			0.23	0.21	0.21	
4. Tipo de alumbrado	0.28			0.31	0.25	0.20	0.32
Bienes							
1. bienes per-cápita	-0.45	-0.53	-0.52	-0.51	-0.42	-0.39	-0.44
2. bienes focalizados per-cápita	-0.43	-0.53	-0.52	-0.49	-0.40	-0.38	-0.43
3. bienes en el hogar	-0.44				-0.42	-0.40	-0.43
4. Licuadora							
5. Refrigeradora/congeladora							
6. Teléfono fijo						-0.19	
7. Lavadora						-0.15	
8. Computadora						-0.22	
9. Internet						-0.14	
10. Cable						-0.21	
Población							
1. Años de educación de mayores de edad		-0.18	-0.31	-0.31	-0.30		
2. Idioma del jefe del hogar			-0.29		-0.22		
3. Ocupación del jefe del hogar							
4. Estado Civil del jefe del hogar							

Fuente: ENAHO 2007-2015. Elaboración: DO-DGGU



Anexo 2

Esquema de los pasos involucrados en la metodología de clasificación socioeconómica 2016.



*Jefe de hogar o cónyuge. **Cualquier integrante.



