

ALGORITMO PARA LA CONSTRUCCIÓN DE LÍNEAS DE POBREZA Y POBREZA EXTREMA A PARTIR DE LA ENCUESTA NACIONAL DE INGRESOS Y GASTOS – ENIG 2006/2007¹

A continuación se realiza una descripción de la información utilizada y del procedimiento empleado para el cálculo de las líneas de pobreza y pobreza extrema. Primero la descripción de los datos utilizados, segundo la metodología definida por la MESEP para la construcción de la línea de pobreza y pobreza extrema, y por último se describe brevemente cada una de las bases de datos necesarias para el cálculo de las líneas de pobreza y pobreza extrema.

1. Fuentes de información

La Fuente de información para el cálculo de la línea es la Encuesta Nacional de Ingresos y Gastos 2006-2007 sobre la cual se realizaron los distintos procedimientos definidos por la MESEP.

1.1 La Encuesta Nacional de Ingresos y Gastos (ENIG) 2006-2007

La nueva línea de pobreza definida por la MESEP está construida a partir de la información más reciente con que cuenta el país para medir detalladamente (a nivel de artículo) los gastos de los hogares, y establecer sus hábitos de consumo: la Encuesta Nacional de Ingresos y Gastos (ENIG) 2006-2007. La desagregación que se utilizó para calcular las líneas fue urbano y rural, dado que dicha desagregación fue considerada por la MESEP como aquella que permitía contar con un número de observaciones suficientes de manera tal que se tuviera confiabilidad en los resultados además de ser el resultado de la desagregación óptima en el sentido en que garantizaba la consistencia, tanto interna como externa, de la encuesta.

Esta encuesta se realizó entre octubre de 2006 y septiembre de 2007, con el objeto de obtener información actualizada y detallada sobre las fuentes de ingreso y la distribución del gasto de los hogares. La información obtenida permite hacer la actualización del Índice de Precios al Consumidor (IPC), conformar la cuenta de hogares en el Sistema de Cuentas Nacionales y además, ser un insumo fundamental para la construcción de las canastas básicas de alimentos con las cuales se calcula las líneas de pobreza y pobreza extrema.

La información de la ENIG tiene dos niveles de presentación: el primero es en el que se ofrece la información recolectada en los distintos formularios aplicados sin ningún tipo de procesamiento, el segundo nivel de presentación ofrece las bases de datos agrupadas en función de distintas necesidades de los investigadores y agregando los gastos por artículo

¹ Documento elaborado por Silvia Botello y Daniel Valderrama, miembros del equipo técnico del DANE que participó en la Misión para el Empalme de las Series de Empleo Pobreza y Desigualdad - MESEP, con base en la metodología definida por esta Misión.

para cada hogar. Para efectos del cálculo de las líneas de pobreza y pobreza extrema se utilizó el segundo nivel de presentación de las bases, más específicamente el archivo llamado conformado 8, que es resultado de la combinación de los conformados 5 y 7:

El conformado 5 contiene la información del gasto por alimento y su cantidad en gramos o centímetros cúbicos mensualizado, obtenido del capítulo D de los formularios 2 tanto urbano como rural. Así mismo incluye el gasto mensual en comidas por fuera del hogar y el gasto en alimentos del formulario de gastos menos frecuentes y de los perceptores de ingreso (DANE, 2009).

El conformado 7 contiene la información del gasto reportado para cada uno de los grupos de alimentos del capítulo C² del formulario 2, urbano y rural. Dado que la información de gastos de este capítulo a diferencia del resto de la encuesta se encuentra agregada por grupos de alimentos, se decidió desagregar al nivel de cada alimento por medio de un procedimiento de extrapolación que consistió en descomponer el gasto agregado reportado por cada hogar en el capítulo C utilizando la distribución promedio del gasto de los hogares que pertenecen a la misma ciudad y decil de ingreso³. La distribución del gasto observada a nivel de alimento es obtenida a partir del capítulo de gastos diarios o semanales registrado por los hogares en el capítulo D de los formularios 2, urbano y rural.

En síntesis, el conformado 8 contiene la información del gasto mensual, cantidad en gramos o centímetros cúbicos para cada alimentos obtenido de la información contenida en los capítulos C y D de los formularios 2, urbano y rural, el formulario 3 de gastos menos frecuentes y el formulario de los perceptores de ingresos (DANE, 2009).

1.2 Redefiniciones de la MESEP sobre el gasto corriente

El gasto de los hogares es usualmente dividido entre gasto no corriente y gasto corriente, el primero está compuesto por los gastos que los hogares realizan esporádicamente y modifica el patrimonio del hogar o de sus integrantes (INEGI 2010) usualmente definidos como gastos en capital, el segundo corresponde a todos los gastos que usualmente los hogares realizan en cada periodo de referencia con el fin de satisfacer sus necesidades de consumo y los gastos de no consumo.

El gasto de los hogares tenido en cuenta para el cálculo de las líneas de pobreza es el gasto que realizan los hogares regularmente y cuya finalidad no es incrementar el patrimonio, es decir, el gasto corriente. Este se toma tanto en términos monetarios como no monetarios. El gasto corriente monetario se define como el gasto, pagado de contado o a crédito, realizado en la compra de bienes y servicios que tienen por objeto satisfacer las necesidades de consumo del hogar, así mismo, incorpora gastos no necesariamente destinados al consumo como es el caso de las transferencias corrientes entre hogares y los impuestos. El gasto corriente no monetario incluye el valor de alimentos, productos y servicios obtenidos de un negocio propio (auto-suministro), producidos en la finca o vivienda (autoconsumo) y los obtenidos como pago en especie.

² Los grupos de alimentos del capítulo C son 25 grupos que agregan distintos alimentos según similitudes en cuanto a su composición.

³ Ver Documento *Resumen de la metodología aplicada para la extrapolación de los gastos del capítulo c de la encuesta de ingresos y gastos 2006-2007*, DANE.

Para calcular el gasto corriente utilizado para el cálculo de las líneas de pobreza la MESEP decidió tener en cuenta tres formas de adquisición:

- **Adquisición:** Se refiere al valor de los bienes y servicios de consumo recibidos por los hogares durante un período determinado, sin tener en cuenta si han sido pagados y tampoco si son consumidos en el período que se adquieren o en períodos siguientes.
- **Pago:** Se refiere a las cancelaciones monetarias o en especie realizadas en el período de referencia sin tener en cuenta si se recibieron o se recibirán y tampoco el período en que serán utilizados y/o consumidos.
- **Consumo:** Se refiere al valor de los bienes y servicios usados o utilizados durante cierto periodo, independientemente de cuándo fueron adquiridos o pagados. Dichos bienes y servicios se “extinguen” al no ser consumidos. Con respecto a los bienes de consumo durable, se debe medir su “uso”.

En las tres formas de adquisición se sumaron: i) gastos de bolsillo en alimentos, transporte, comunicaciones y recreación; y ii) los gastos menos frecuentes con las siguientes excepciones: cuota amortización de vivienda diferente a la que habita, compra de vivienda, la ampliación o subdivisión, compra de terrenos, lotes, locales, fincas, parcelas y similares, pago de impuestos de las viviendas que ocupa el hogar, compra de vehículos nuevos y usados para uso particular, cuotas mensuales finca raíz o vehículos, gastos de matrícula para vehículo nuevo y de traspaso para vehículos usados, certificado de gases y otras revisiones ordenadas por la ley, pago de SOAT, impuestos del vehículo, impuesto de renta y otros impuestos.

Además, dentro de la definición de gasto corriente se incluyeron los rubros de gastos de reparación y mantenimiento de muebles accesorios y materiales para pisos, compra de libros, arriendo imputado, materiales y mano de obra para el mantenimiento de vivienda, teniendo en cuenta los criterios de durabilidad del bien, valor y frecuencia de compra⁴, y también el pago de tarjetas de crédito

Finalmente, se descontó el valor por concepto de salud, pensión, ICA, IVA, retefuente, diario oficial, descuento solidaridad pensional, deducciones hechas a los ocupados tanto asalariados como independientes.

2. El cálculo de las líneas de pobreza y pobreza extrema

Utilizando la Encuesta Nacional de Ingresos y Gastos ENIG (2006-2007) procesada por el DANE en el conformado 8 y con las modificaciones hechas por la MESEP a la definición del gasto corriente de los hogares se procede a calcular las líneas de pobreza y pobreza extrema. Para el cálculo de las líneas se realizan tres grandes procedimientos:

- 2.1 El cálculo del deflactor espacial de precios.
- 2.2 La estimación de las canastas de alimentos.

⁴ No se incluyeron gastos en muebles, bicicleta, bienes de consumo durable, ampliación de vivienda (inversión), ventiladores y abanicos, equipo y accesorios para camping, enciclopedias, obras y colecciones, radio para vehículo.

2.3 El cálculo del valor de las líneas.

2.1 Deflactor Espacial de Precios (DEP)

Dado que el poder adquisitivo de los hogares cambia entre espacios geográficos, al momento de seleccionar una población homogénea en función de su nivel de gasto corriente debe observarse el gasto real en lugar del gasto nominal, es decir el gasto en términos de la capacidad de compra del mismo. Por tanto se deflacta el gasto corriente nominal usando un deflactor espacial de precios (DEP), el cual se calcula de la siguiente forma:

- Tomar la base ENIG_alimentos y excluir las comidas por fuera del hogar, es decir, se deja únicamente los artículos correspondientes a alimentos y bebidas no alcohólicas (Capítulo 1 de la nomenclatura de CCIF⁵)
- Eliminar los registros que presentan observaciones inconsistentes: aquellas que tienen 0 o sin información en variables como cantidad consumida (cantot), valor del gasto (valtot).
- Calcular el logaritmo natural del valor unitario, este último se define como el valor de compra sobre la cantidad de compra (solo se calcula el logaritmo natural para los valores unitarios mayores a cero), con base en este cálculo se procede a excluir las observaciones atípicas, definidas como aquellas que para cada dominio (urbano y rural) y artículo se encuentren por fuera de un intervalo de confianza de 2.5 desviaciones estándar⁶.
- Seleccionar los artículos que se incluirán en el cálculo del deflactor espacial de precios, estos serán todo artículo que para cada dominio (urbano y rural) cumpla con las siguientes condiciones:
 - Que sea consumido por 30 hogares por cada dominio
 - Que los alimentos estén tanto en la zona urbana como rural.
- Calcular sobre la base de datos que cuenta únicamente con los artículos que cumplen la condición anterior, la mediana del valor unitario de cada artículo por dominio considerando el factor de expansión de personas del hogar
- Calcular la participación que tiene el gasto dentro del gasto total para cada artículo por dominio (promedio de la participación del gasto de cada artículo dentro del gasto total que se observa en los hogares al interior de cada dominio, este procedimiento es definido en la literatura como ponderador democrático⁷) considerando el factor de expansión de personas del hogar

⁵ Clasificación del Consumo Individual por finalidades.

⁶ Para este cálculo se considera el factor de expansión del hogar

⁷ Los hogares que no reportan consumo en un artículo implica que la participación de este en el gasto es cero.

- Finalmente se obtiene una base de datos que para artículo y dominio (urbano rural) cuenta con información sobre los precios y la participación que el artículo tiene en el gasto total en alimentos, sobre esta base de datos se aplica la fórmula del índice de precios de Tornqvist, considerando como unidad espacial de referencia la zona urbana:

$$DEP_{ru} = \prod_i^m \left(\frac{p_{ir}}{p_{iu}} \right)^{\frac{1}{2} \left[\frac{p_{iu}q_{iu}}{\sum_{i=1}^m (p_{iu}q_{iu})} + \frac{p_{ir}q_{ir}}{\sum_{i=1}^m (p_{ir}q_{ir})} \right]}$$

Donde i indexa los 220 alimentos que cumplían con los dos criterios de selección para el cálculo del DEP, p_{ir} es la mediana de la distribución de precios del bien i en la zona rural; p_{iu} es la mediana de la distribución de precios del bien i en la zona urbana; $\frac{p_{iu}q_{iu}}{\sum_{i=1}^m (p_{iu}q_{iu})}$ es el promedio para la zona urbana de la participación que tiene el bien i dentro del gasto total en alimentos y $\frac{p_{ir}q_{ir}}{\sum_{i=1}^m (p_{ir}q_{ir})}$ es el promedio para la zona rural de la participación que tiene el bien i dentro del gasto total en alimentos.

2.2 Población de referencia

La construcción de una línea de pobreza requiere identificar los patrones de consumo y precios a los que se enfrenta la población que esta alrededor de lo que se considera el umbral de la pobreza, es decir no se toma los más pobres ni los más ricos, se toman los que están alrededor de la línea de pobreza; en términos operativos, se toma la población cuyos percentiles de gasto corriente se encuentra alrededor de la tasa de incidencia de la pobreza que se obtenía con la anterior metodología, es decir el percentil 45. Así la población que se encuentra alrededor de este percentil de acuerdo a su gasto en términos reales es sobre la cual se definen los patrones de consumo necesarios para el cálculo de la línea, por tanto los pasos para esto son:

- Tomar la base de datos a nivel de hogares ENIG_gasto_corriente y deflactar el gasto corriente per-cápita de cada hogar usando el deflactor espacial de precios de Tornqvist obtenido en el cálculo anterior.
- Calcular el percentil al que pertenece cada hogar y dejar solamente los hogares que se encuentran entre los percentiles 30 y 59, este intervalo tiene un rango de 30 percentiles alrededor del percentil 45, para este cálculo se considera el factor de expansión de las personas de la unidad de gasto.
- Pegar a los hogares de la población de referencia la información del consumo de alimentos que se encuentra en la base ENIG_alimentos, para pegar ambas bases se usa el identificador de vivienda y hogar, de ahora en adelante se trabaja únicamente con los hogares de la población de referencia.
- Depurar la base de valores inconsistentes: Se borran los alimentos que no estén identificados (que la variable de artículo no tenga información), los que tengan

valor de compra cero o sin información, y los alimentos para los que no se cuente con información sobre el aporte calórico.

- Seleccionar los alimentos que van en la canasta de cada dominio, zona urbana y rural por separado. Para que un artículo sea seleccionado dentro de la canasta corta de alimentos de un dominio debe cumplir con al menos uno de los siguientes tres criterios, por tanto se debe calcular una variable por criterio que defina si el hogar cumple o no con este y si alguna de las variables indica cumplimiento se deja dicho artículo (para observar el cálculo de los criterios se usa el factor de expansión del hogar) , los 3 criterios mencionados son:
 - Que el artículo sea consumido por lo menos por un 30% de los hogares de cada dominio.
 - Que el artículo represente más del 1% del gasto en alimentos de cada dominio
 - Que el artículo represente más del 1% del consumo calórico de cada dominio.
- Calcular el coeficiente de Engel y con este el coeficiente de Orshansky para cada dominio. El coeficiente de Engel es definido como la participación que tiene el gasto en alimentos (incluyendo tanto las comidas por fuera del hogar como los alimentos que se compran para el hogar) sobre el gasto corriente total. El coeficiente de Orshansky es definido como el inverso del coeficiente de Engel.

Al final de este procedimiento tendremos una base que contenga únicamente los hogares de la población de referencia y el consumo que estos realizan en los alimentos que hacen parte de la canasta por cumplir con al menos uno de los criterios, además se tendrá el cálculo del coeficiente de Orshansky asociado a cada dominio.

2.3 Valor de la Canasta

Teniendo en cuenta la canasta de alimentos de cada dominio y los requerimientos calóricos diarios de la FAO/OMS/ONU (2001), podemos calcular el costo de la canasta básica de alimentos que satisface dichos criterios calóricos, que corresponde a la línea de indigencia. Para esto utilizamos en cada dominio las cantidades que deberían ser consumidas en promedio y la mediana de los valores unitarios que paga cada hogar de la población de referencia. Para esto partimos de la base con la cual se terminó el procedimiento anterior es decir con una base de datos que solo contiene los hogares de la población de referencia y su consumo en los alimentos que pertenecen a la canasta corta:

- Calcular las calorías consumidas efectivamente por la población de referencia en cada dominio al considerar únicamente los alimentos seleccionados en la canasta. Para esto se multiplica las cantidades consumidas en gramos o centímetros cúbicos por el porcentaje comestible de dicho producto y por su contenido calórico, una vez obtenido este valor se suman las calorías totales que son consumidas por

dominio y se dividen por el número de personas de cada dominio⁸, con esto obtenemos el promedio del consumo calórico per-cápita mensual, para volverlo en unidades diarias, igual que el dato del requerimiento calórico per cápita diario , procedemos a dividirlo por 30.

- Calcular el factor de ajuste para cada dominio, este es el cociente entre el requerimiento calórico per-cápita diario de cada dominio⁹ y el cálculo del consumo calórico per-cápita diario efectivo hecho en el paso anterior.
- Ajustar las cantidades de consumo multiplicando las cantidades consumidas por el factor de ajuste
- Calcular la mediana del valor unitario de cada alimento de la canasta para cada dominio (usando el factor de expansión de los hogares)
- Para cada dominio se calcula la sumatoria del producto entre las medianas de los valores unitarios y las cantidades ajustadas de cada artículo dando como resultado el costo per-cápita mensual de la canasta de alimentos, definido como la línea de pobreza extrema.
- Calcular el factor de ajuste del coeficiente de Orshansky como la razón entre el coeficiente de Orshansky rural y coeficiente de Orshansky urbano
- Multiplicar la línea de pobreza extrema urbana por el coeficiente de Orshansky exógeno adoptado por la MESEP de 2.4 y así se obtiene la línea de pobreza urbana y la línea de pobreza rural se calcula multiplicando la línea de pobreza extrema por 1.74 (El producto entre el coeficiente exógeno (2.4) y el factor de ajuste del coeficiente de Orshansky calculado en el paso anterior).
- Las líneas obtenidas corresponden a las líneas de pobreza y pobreza extrema para el dominio urbano y rural de Marzo de 2007, a partir de entonces la línea de pobreza extrema se actualiza usando el IPC de alimentos de ingresos bajos por ciudades y la línea de pobreza usando el IPC total de ingresos bajos. Estos están disponibles en el sistema de consulta de la página web del DANE.

⁸ Estrictamente serían las personas de la población de referencia pues la base con la que trabajamos solo se encuentra la población de referencia.

⁹ Esta es el requerimiento calórico per cápita diario promedio del dominio, basado en la estructura (género y edad) de cada zona urbano y rural con base en la información de la FAO/OMS/ONU (2001)