



INSTITUTO NACIONAL DE ESTADÍSTICA Y CENSOS

***ENCUESTA DE HOGARES DE PROPOSITOS MULTIPLES
METODOLOGIA DE EDUCACION***

Aspectos metodológicos

A. Objetivos

La Encuesta de Hogares de Propósitos Múltiples, que se realiza en julio de cada año, tiene como objetivos principales los siguientes: (a) mantener un flujo continuo de estadísticas relacionadas con la fuerza de trabajo, el empleo, el desempleo, el subempleo y los ingresos, y de otras variables socioeconómicas necesarias para el establecimiento de políticas y para la formulación de planes orientados al desarrollo económico y social del país, así como para la evaluación de sus efectos; (b) proveer información periódica, sistemática y oportuna en los períodos intercensales referente, entre otras, a las variables mencionadas; y (c) servir de fuente de información a instituciones gubernamentales, universitarias o de investigación, interesadas en temas relativos a la población y el empleo.

B. Conceptos y definiciones

1. Grado de urbanización

Es el grado de urbanización asignado a cada uno de los segmentos censales, de acuerdo con los siguientes criterios:

Urbana: se clasifican bajo esta categoría los centros administrativos de los cantones del país; esto es, parte o todo el distrito primero de cada cantón, centros de otros distritos y otras áreas adyacentes que cumplen con el criterio físico y funcional que toma en cuenta elementos tangibles como: cuadrantes, calles, aceras, luz eléctrica y servicios urbanos.

Periferia urbana: se clasifican bajo esta categoría aquellas zonas que pertenecen a las áreas localizadas entre el límite del cuadrante urbano y la poligonal envolvente del área urbana.

La poligonal envolvente consiste en una línea imaginaria que encierra, tanto el cuadrante o casco urbano como las áreas adyacentes al mismo, que aunque no se encuentran dentro del cuadrante urbano, presentan características similares a las de estas zonas y se les considera como áreas en transición de rural a urbano.

Rural concentrado: se clasifican bajo esta categoría todos aquellos centros poblados no ubicados en las categorías anteriores y que reúnen las siguientes características:

- El suelo está ocupado predominantemente por actividades no agropecuarias.
- Tienen 50 o más viviendas agrupadas o contiguas, en general la distancia entre ellas no es mayor a los 20 metros.
- Disponen de algún servicio de infraestructura como electricidad domiciliaria, agua potable o teléfono.
- Cuentan con algunos servicios como escuela, iglesia, centro de salud, puestos de salud y guardia rural.

Rural disperso: todas las áreas que no pertenecen a alguna de las clasificaciones anteriores.

2. Zona

Responde a la clasificación de “urbana” o “rural” que bajo ciertos criterios se asigna a los segmentos censales de acuerdo a su grado de urbanización.

En encuestas anteriores al año 2000, se consideraban “Urbanas” las zonas con grado de urbanización urbana; y “Rurales” las que tenían la calificación de periferia urbana, rural concentrado y rural disperso. A partir del Censo del 2000, y considerando que en la periferia urbana predominan características que se asemejan más a la de una zona urbana, se decidió unir esta categoría con la “Urbana”. Por lo anterior, y para mantener comparabilidad en las clasificaciones del Censo y de la Encuesta, a partir de la Encuesta del año 2000, se aplica la siguiente clasificación en la tabulación de los datos:

Urbana: se incluye bajo esta categoría a aquellos segmentos clasificados con grado de urbanización “urbana” y “periferia urbana”.

Rural: se incluye bajo esta categoría a aquellos segmentos clasificados con grado de urbanización “rural concentrado” y “rural disperso”.

C. Características del Diseño Muestral

1. Población de estudio

La población de estudio de la Encuesta de Hogares de Propósitos Múltiples está definida como el conjunto de todas las viviendas particulares existentes en el país y de sus ocupantes, que son residentes habituales en esas viviendas. Se excluye del estudio a la población residente en las viviendas colectivas (cárceles, conventos, asilos, residencias colectivas para estudiantes y trabajadores, hospitales y hoteles).

2. Marco muestral de viviendas

Los marcos muestrales de viviendas son los instrumentos utilizados para la selección de muestras de viviendas. Para seleccionar las muestras de las EHPM, entre 1987 y 1998 se empleó un marco muestral de viviendas construido en 1986 con la cartografía y la información de los Censos Nacionales de Población y Vivienda de 1984 (MMV-86).

Considerando que el crecimiento de la población desde esa fecha y los cambios ocurridos en su distribución geográfica, caracterizada especialmente por un proceso de creciente urbanización del país, había modificado la situación existente y por consiguiente provocado la desactualización del MMV-86, en 1997, la entonces Dirección General de Estadística y Censos inició acciones orientadas a la elaboración de un nuevo marco muestral que contemplara apropiadamente el crecimiento en el número de viviendas y los cambios en la condición urbano-rural de las áreas o segmentos que lo componen.

Como parte de este proceso se construyó un nuevo marco (MMV-98), con base en la actualización cartográfica realizada en el período 1993-1997, actividad que formó parte de los

preparativos del censo que finalmente se llevó a cabo en julio del año 2000^{1/}. Este MMV presenta una nueva delimitación de segmentos, incluye un conteo del número de viviendas dentro de cada segmento y algunas variables que pueden ser utilizadas como criterios de estratificación: la región de planificación^{2/}, el grado de urbanización (urbana, periferia urbana, rural concentrado y rural disperso), y un indicador del nivel socioeconómico (bajo, medio - bajo, medio, medio - alto, alto), construido a partir de una calificación visual de las viviendas y su entorno que se realizó para cada uno de los segmentos de la zona urbana y de la periferia urbana del país^{3/}.

El MMV de 1998 quedó constituido por 16106 segmentos y 826 521 viviendas con la distribución por Regiones de Planificación que se aprecia en el cuadro N° 1. Con este nuevo marco se seleccionaron las muestras utilizadas en la encuesta de 1999 y las subsiguientes.

Actualmente, se trabaja en la elaboración de un nuevo marco muestral que toma como referencia la cartografía y la información del Censo 2000, y que se espera aplicar en próximas encuestas, una vez concluida la preparación y el diseño de una nueva muestra.

La regionalización del país utilizada para esta Encuesta se ilustra en el mapa de la siguiente página.

3. Dominios de estudio

Los dominios de estudio son áreas geográficas de desagregación que se consideran en el diseño de la muestra para tener niveles de precisión aceptables (conocidos a priori).

En el caso de la EHPM la muestra se diseñó para tener niveles de precisión aceptables para las principales variables que investiga la encuesta en los siguientes dominios de estudio:

- Región Central urbana
- Región Central rural
- Resto del país urbano
- Resto del país rural

¹ Lo deseable habría sido elaborar el nuevo Marco Muestral de Viviendas con los datos de un nuevo censo, no obstante, ante la incertidumbre respecto a la fecha en que finalmente se iba a realizar, se decidió elaborar el MMV-98 con base en las actualizaciones cartográficas realizadas entre 1993 y 1997.

² Según Decreto N° 160068-PLAN del 15 de febrero de 1985, La Gaceta N° 59 del 26 de marzo de 1985.

³ Asignación del Nivel Socioeconómico. Proceso de Muestreo, 1998.

CUADRO 1

Número de segmentos y viviendas del Marco Muestral de Viviendas de 1998 por zona, según región de planificación

Región de planificación	Total		Zona			
	Segmentos	Viviendas	Urbana		Rural	
			Segmentos	Viviendas	Segmentos	Viviendas
Costa Rica	16 106	826 521	7 781	463 594	8 325	362 927
Central	9 669	525 076	6 105	363 854	3 564	161 222
Chorotega	1 419	67 722	387	23 593	1 032	44 129
Pacífico Central	992	50 061	461	26 359	531	23 702
Brunca	1 452	64 715	273	16 653	1 179	48 062
Huetar Atlántica	1 669	78 349	439	25 936	1 230	52 413
Huetar Norte	905	40 598	116	7 199	789	33 399

Regionalización de Costa Rica^{1/}



Sin embargo, dependiendo del comportamiento de algunas características en la población y de los resultados de la encuesta (a posteriori), la muestra permite dar estimaciones con niveles de precisión aceptables para otras desagregaciones geográficas como lo son las regiones de planificación.

4. Diseño muestral

El diseño muestral corresponde a un diseño probabilístico de áreas, estratificado y bietápico. Es de áreas debido a que las probabilidades de selección están asociadas a los segmentos censales, los cuales son áreas geográficas debidamente delimitadas; es estratificado porque para la distribución y selección de la muestra se definieron doce estratos de interés (conformados por cada región de planificación dividida por zona urbana y rural) con la finalidad de tener una mejor

representatividad de estas áreas y aumentar así la eficiencia relativa del diseño; es bietápico ya que en una primera etapa se seleccionan segmentos censales o Unidades Primarias de Muestreo (UPM), y en una segunda etapa se seleccionan viviendas o Unidades Secundarias de Muestreo (USM) dentro de las UPM seleccionadas en la primera etapa.

El diseño muestral establece que la muestra sea autoponderada dentro de cada estrato. Por ello, como los segmentos se seleccionaron con igual probabilidad, para mantener la autoponderación se toma una fracción fija de viviendas en la segunda etapa de selección: un cuarto (1/4) en los segmentos urbanos y un tercio (1/3) en los segmentos rurales.

5. Tamaño y distribución de la muestra

El cálculo del tamaño de muestra se hace bajo el supuesto de muestreo aleatorio simple con un ajuste por efecto de diseño por tratarse de una muestra compleja^{11/}. El cálculo del tamaño de muestra y su distribución se efectuó de la siguiente manera:

- a. Se procedió a calcular el tamaño de muestra necesario para obtener estimaciones de una tasa de desempleo del 5% , con un margen de error del 1%, un nivel de confianza del 95% y un efecto del diseño de 2,45.
- b. Dado que se establecieron cuatro dominios de estudio (Región Central Urbana, Región Central Rural, Resto del País Urbano y Resto del País Rural), el tamaño de muestra para el total del país corresponde al cálculo anterior multiplicado por cuatro.
- c. Considerando el interés del país en la obtención de información sobre el trabajo infantil por medio de la aplicación de módulos especiales, se asignó o distribuyó la muestra nacional por zona urbana y rural de acuerdo a la variabilidad de la tasa de participación económica de la población de 5 a 17 años, estimada a partir del módulo de trabajo infantil aplicado en el año 1995.
- d. La muestra resultante en cada zona, se distribuyó por región, de acuerdo a la variabilidad relativa del número de desempleados, por tanto no es una muestra proporcional al marco.
- e. De esta forma se obtuvo un tamaño de muestra inicial de 10 890 viviendas, el cual ha aumentado a 15 242 en el año 2009, por el crecimiento en el número de viviendas. La muestra se distribuye de la siguiente forma:

CUADRO 2

Número de segmentos y viviendas seleccionadas por zona, según región de planificación
JULIO 2009

Región de planificación	Total		Zona			
	Segmentos	Viviendas	Urbana		Rural	
			Segmentos	Viviendas	Segmentos	Viviendas
Costa Rica	726	15 242	322	5 954	404	9 288
Central	334	6 771	200	3 660	134	3 111
Chorotega	82	1 964	28	575	54	1 389
Pacífico Central	74	1 452	32	488	42	964
Brunca	92	1 866	24	449	68	1 417
Huefar Atlántica	92	1 927	26	396	66	1 531
Huefar Norte	52	1 262	12	386	40	876

11/ Una muestra se considera compleja cuando utiliza más de una etapa de selección, como es el caso de la EHPM.

D. Sobre las estimaciones de los resultados

1. Procedimiento para la estimación de los valores poblacionales

El procedimiento de estimación de la Encuesta consiste en multiplicar el resultado de la muestra por un factor de expansión. Este procedimiento se aplica para cada uno de los doce estratos de selección (las seis regiones por zona urbana y rural) de la siguiente forma:

- a) **Factor de expansión inicial:** se calcula el factor de expansión como el inverso de la fracción final de muestreo para cada estrato. La fracción final de muestreo es el producto de la fracción de muestreo en la primera etapa y la fracción de muestreo en la segunda etapa de selección.
- b) **Ajuste por crecimiento:** se calcula un ajuste por crecimiento de viviendas que consiste en la razón del número de viviendas en el marco actualizado por el crecimiento observado en los segmentos o UPMs de la muestra (número inicial más las nuevas viviendas) entre el número original de viviendas en el marco.
- c) **Ajuste por no respuesta:** se calcula un ajuste por no respuesta que consiste en la razón del número total de viviendas seleccionadas sujetas a entrevista entre el número de estas viviendas que efectivamente respondieron la entrevista.
- d) **Ajuste de población:** se hace con el fin de corregir la subestimación mencionada en el apartado D.I. Este ajuste consiste en la razón de la población proyectada con base en la información censal y la estimación de la encuesta.
- e) **Factor de expansión final:** el factor de expansión final se obtiene como el producto del factor de expansión inicial, el ajuste de crecimiento, el ajuste por no respuesta y el ajuste de población.

Para obtener la estimación de población para las diferentes variables que investiga la

encuesta, se multiplica la variable o característica de cada persona por su correspondiente factor de expansión, y el resultado se agrega para cada ámbito geográfico: región, zona urbana o rural y nacional.

En el caso de las estimaciones sobre los hogares, el factor de expansión que se aplica a un hogar es el factor de expansión correspondiente al jefe de ese hogar. Las características del hogar se multiplican por este factor de expansión y el resultado se agrega de igual forma para cada ámbito geográfico.

2. Confiabilidad de los datos

La muestra utilizada en la encuesta es una de todas las muestras posibles del mismo tamaño que podrían haberse seleccionado utilizando el mismo diseño muestral; cada muestra proporciona una estimación del valor poblacional que se desea conocer. El error de muestreo representa una medida de la dispersión o variabilidad de las estimaciones de todas las muestras posibles con respecto a ese valor poblacional que se desea estimar y es de hecho una medida de la precisión de las estimaciones.

Debido a que se selecciona sólo una muestra de todas las muestras posibles, el error de muestreo no es posible calcularlo directamente sino que se estima por medio del error estándar que es una medida del nivel de precisión de las estimaciones de la encuesta. En este sentido, se recomienda consultar el apartado IV sobre los resultados de los errores de muestreo tales como coeficiente de variación, error estándar y límites de confianza los cuales han sido calculados para las principales variables.

Para analizar las estimaciones, es importante tomar en consideración que las celdas con valores pequeños representan características poco frecuentes en la población y, por tanto, los resultados deben ser tratados con cuidado.

E. Sobre las variables utilizadas

1. Variables para estimaciones de Educación

Para obtener los indicadores referentes al tema de educación se trabaja con variables de edad, sexo, último grado o año aprobado, asistencia a centros educativos, entre otras.

Para la población de 15 años o más se obtienen distribuciones por nivel educativo, títulos obtenidos, estudios en educación no formal y principales centros de estudio. Además estos indicadores se tienen disponibles para las regiones de planificación y zonas geográficas.