

ASPECTOS METODOLÓGICOS DE LA ENCUESTA PERMANENTE DE HOGARES (EPH) 2016

1. OBJETIVO

El objetivo principal de la Encuesta Permanente de Hogares 2016 es generar indicadores relacionados con el empleo, el desempleo, los ingresos y otras características sociales y económicas, que permitan conocer la evolución del bienestar de la población paraguaya.

2. ÁMBITO GEOGRÁFICO

La muestra total de la EPH 2016 incluye los 17 departamentos del país más Asunción.

3. ÁMBITO POBLACIONAL

La investigación va dirigida a la población que reside habitual o permanentemente en viviendas particulares.

Se excluye de la investigación a la población residente en las viviendas colectivas. Esta categoría comprende: los hoteles, pensiones y otras casas de huéspedes; aunque, sí se incluyen a las familias que, formando un grupo independiente residen dentro de estos establecimientos, como puede ocurrir con los directores de los centros, conserjes, porteros, etc.

4. PERÍODO DE LEVANTAMIENTO DE LOS DATOS

El levantamiento de los datos corresponde al período de octubre a diciembre de 2016.

5. CARACTERÍSTICAS DEL DISEÑO MUESTRAL

5.1. MARCO MUESTRAL

El marco utilizado para la selección de la muestra de la región oriental no indígena está constituido por “segmentos”, conformados a partir del Censo Nacional de Población y Viviendas 2012.

Para la muestra, los segmentos utilizados como Unidades Primarias de Muestreo (UPM) es, el Área de Trabajo del Censista (ATC), con un promedio de 60 viviendas en el ámbito urbano y 45 viviendas en el ámbito rural.

Para los departamentos de la región occidental no indígena el marco está constituido también por “segmentos”, conformados a partir del Censo Nacional de Población y Viviendas 2012.

Para la muestra, los segmentos utilizados como Unidades Primarias de Muestreo (UPM) es, el Área de Trabajo del Censista (ATC), con un promedio de 60 viviendas en el ámbito urbano y 25 viviendas en el ámbito rural.

Para la selección de hogares que pertenecen a comunidades indígenas las Unidades Primarias de Muestreo (UPM) están compuestas por comunidades indígenas.

5.2. NIVELES DE ESTIMACIÓN

Los niveles de desagregación para los cuales el diseño muestral permite obtener estimaciones con un nivel de confianza conocido son los siguientes:

- Total país
- Total urbano
- Total rural
- Asunción
- Todos los departamentos

5.3. ESTRATIFICACIÓN GEOGRÁFICA

Las variables de estratificación son, el departamento geográfico y el área de residencia. La EPH 2016 tendrá un total de 35 estratos. Éstos comprenderán Asunción, más los 17 departamentos, divididos cada uno en estratos urbano y rural.

5.4. TIPO DE MUESTREO

El diseño muestral corresponde a un diseño probabilístico por conglomerados con probabilidad proporcional al tamaño, bietápico y estratificado en la primera etapa.

Es bietápico, ya que en una primera etapa se seleccionan dentro de cada estrato los segmentos censales o Unidades Primarias de Muestreo (UPM) y, en una segunda etapa, se seleccionan las viviendas o Unidades Secundarias de Muestreo (USM).

5.4.1. Selección de viviendas

Primera etapa de selección: dentro de cada estrato, se selecciona un número determinado de UPMs en forma sistemática, con arranque aleatorio y con probabilidad proporcional al tamaño.

Segunda etapa de selección (Submuestreo): dentro de cada UPM escogida en la etapa anterior, en el marco no indígena se seleccionan aleatoriamente un total fijo de 12 viviendas y en el marco indígena se seleccionan aleatoriamente un total fijo de 10 viviendas. En ambas ocasiones la selección fue sin reemplazo.

Las viviendas se consideran como conglomerados; en ellas se investigan a todas las personas que habitualmente residen.

5.5. TAMAÑO Y DISTRIBUCIÓN DE LA MUESTRA

El tamaño de la muestra total de la EPH 2016 fue de 13.056 viviendas, distribuidas en todos los departamentos del país de manera a lograr su correspondiente representatividad. La determinación de dicho tamaño muestral contempla una tasa de no respuesta del 10%.

La muestra fue distribuida en estratos de la siguiente forma:



DISTRIBUCIÓN DE LA MUESTRA POR ESTRATOS, EPH 2016

Muestra del marco no indígena

DEPARTAMENTO	TOTAL		URBANA		RURAL	
	UPMs	VIVIENDAS	UPMs	VIVIENDAS	UPMs	VIVIENDAS
Asunción	80	960	80	960	-	-
Concepción	37	444	16	192	21	252
San Pedro	52	624	12	144	40	480
Cordillera	72	864	28	336	44	528
Guairá	49	588	21	252	28	336
Caaguazú	61	732	26	312	35	420
Caazapá	50	600	11	132	39	468
Itapúa	77	924	35	420	42	504
Misiones	50	600	28	336	22	264
Paraguarí	40	480	13	156	27	324
Alto Paraná	118	1.416	87	1.044	31	372
Central	98	1.176	84	1.008	14	168
Ñeembucú	39	468	24	288	15	180
Amambay	50	600	36	432	14	168
Canindeyú	56	672	24	288	32	384
Pdte. Hayes	22	264	8	96	14	168
Boquerón	34	408	17	204	17	204
Alto Paraguay	18	216	9	108	9	108
Total	1.003	12.036	559	6.708	444	5.328

Muestra del marco indígena

DEPARTAMENTO	TOTAL	
	UPMs	VIVIENDAS
Concepción	3	30
San Pedro	3	30
Guairá	1	10
Caaguazú	8	80
Caazapá	3	30
Itapúa	2	20
Alto Paraná	6	60
Central	2	20
Amambay	10	100
Canindeyú	13	130
Presidente Hayes	24	240
Boquerón	18	180
Alto Paraguay	9	90
Total	102	1.020

6. CALIDAD DE LAS ESTIMACIONES

Las estimaciones de toda encuesta realizada por muestreo vienen afectadas por dos tipos de errores:

- a) **Errores de muestreo:** los datos de una muestra están sujetos a los errores de muestreo, los cuales se presentan debido a que la investigación estadística se hace sólo en una parte de la población, para estimar características de todo el conjunto poblacional. En general, a medida que aumenta el tamaño de la muestra disminuye el error de muestreo, aunque también influyen en él las características del diseño y la naturaleza de la variable investigada.
- b) **Errores ajenos al muestreo:** incluyen toda una serie de errores, que se producen en las distintas fases de realización de una encuesta, desde el diseño del cuestionario hasta la publicación de los resultados finales (definiciones deficientes, errores en el marco, falta de respuesta en las unidades informantes, errores de codificación y grabación etc.).

7. ERRORES DE MUESTREO

Los errores de muestreo no se conocen directamente, pero sí es posible estimarlos a partir de los datos de la propia muestra. En tal sentido, mediante el software SPSS utilizado actualmente se obtiene, (ver anexo), las estimaciones del error de muestreo, error relativo o, coeficiente de variación, intervalos de confianza, de las variables más importantes de la Encuesta Permanente de Hogares, del Módulo de Empleo, como son: Tasa de actividad, Tasa de ocupación, Tasa de desempleo, Estatus de pobreza, etc., tanto a nivel nacional como por algunos dominios de estudio.

A fin de interpretar de una manera sencilla estos resultados, seguidamente se presenta el concepto de cada columna de los cuadros.

7.1. Estimación.

Para valores absolutos, son datos muestrales multiplicados por el factor de expansión, para proporción o tasa son valores porcentuales obtenidos mediante las formulas arriba explicadas para cada variable en estudio.

7.2. Error muestral de la estimación. Es el valor absoluto de la raíz cuadrada de la varianza del estimador. Al error de muestreo se le denomina también error estándar y es una medida de la precisión con que una estimación de la muestra se aproxima al promedio de todas las muestras posibles.

La varianza es una medida estadística que permite determinar el grado de dispersión de la variable o el grado de heterogeneidad u homogeneidad de los elementos de un conjunto de elementos en función de los valores que la variable toma en cada uno de los elementos del conjunto.

Una manera sencilla de interpretar el error estándar (e.e.) es en término de intervalo de confianza.

7.3. Coeficiente de Variación (C.V.). El coeficiente de variación (CV) es definido como el error estándar de la estimación dividido por el valor estimado. Se lo conoce como error relativo.

El coeficiente de variación indica más claramente el nivel de precisión de una estimación; en efecto, en las encuestas de hogares la experiencia ha demostrado que estimaciones con un

coeficiente de variación de hasta un 5 por ciento son muy precisas; si el coeficiente de variación llega hasta un 10 por ciento, las estimaciones siguen siendo precisas; un coeficiente de variación con un valor de hasta 20 por ciento es aceptable; y por último, más allá de un 20 por ciento indica que la estimación es poco confiable y, por tanto, se debe utilizar con precaución.

7.4. Intervalo de Confianza al 95%. Es el intervalo estimado que incluye el promedio de las estimaciones de todas las muestras posibles con una asignada probabilidad de confianza. Este intervalo tiene un límite inferior y un límite superior, dentro de los cuales se espera que se encuentre el valor real o poblacional con el nivel de confianza determinado.

Para cualquier estimador “y” de la encuesta, se puede obtener los intervalos de confianza de la manera siguiente:

$$\text{Estimador } y \pm z * e.e.$$

La expresión $z * e.e.$ recibe el nombre de margen de error del estimador y viene dado por el producto entre t y el error estándar de la estimación.

El valor z , es la abscisa de la distribución normal para una determinada probabilidad. Para un nivel de confianza del 90%, el valor de $z = 1,645$; para el 95%, $z = 1,96$, y para 99%, $z = 2,58$.

El nivel de confianza que va a ser usado en el análisis puede ser determinado por el usuario de la información. Habitualmente se toma el 95%.

La amplitud del intervalo de confianza depende principalmente del error estándar; es decir, a mayor valor del error estándar mayor amplitud del intervalo de confianza y, por tanto, menor precisión del valor estimado. Lo esperable es tener una menor amplitud del intervalo de confianza, por tanto, mayor precisión del valor estimado.

8. CUESTIONARIO

El cuestionario utilizado para la Encuesta Permanente de Hogares 2016 consta de 8 secciones bien diferenciadas y son las siguientes:

1. Características de la Población: En esta parte se investiga las características demográficas de las personas que residen habitualmente en el hogar, es decir, su edad, sexo, estado civil, lugar de nacimiento y migración interna. Esta parte permite identificar a los “miembros del hogar” que residen en la vivienda.

2. Características de la Vivienda e Inventario de bienes duraderos: Se recoge información sobre el tipo de vivienda, tipo de material que predomina en su construcción, los servicios con que cuenta, así como los bienes duraderos que tiene el hogar.

3. Tecnologías de la Información y la Comunicación: Se recogerá información sobre acceso a diferentes medios tecnológicos a las personas de 10 y más años de edad.

4. Educación: En esta sección interesa conocer los aspectos relacionados con el nivel educativo, asistencia a centros educativos de todas las personas de 5 años y más. Las estimaciones que surjan de dicho módulo son utilizadas principalmente para relacionarlas con los indicadores del mercado laboral.

5. Salud: Contiene preguntas sobre tenencia de seguro médico, el estado de salud de cada uno de los miembros, la utilización de los servicios en los últimos 90 días en caso de enfermedad o accidente.

6. Empleo e Ingreso laboral: Determina si los miembros del hogar de 10 años y más de edad realizan o no actividades económicas, (investiga trabajo dependiente, independiente no agropecuario y agropecuario). Se investiga la ocupación principal y secundaria. Recoge también información acerca de los ingresos y las fuentes de donde provienen. Además se



recoge información detallada acerca de las remesas que provienen como ayuda familiar del exterior.

7. Actividades Independientes no Agropecuarias del Hogar: Investiga si en el hogar existe algún miembro que tenga un negocio o empresa por su cuenta. En caso afirmativo, se capta información sobre los ingresos y gastos de la empresa o negocio.

8. Actividades Agropecuarias, Forestales y Derivados del Hogar: Investiga la tenencia de tierras y maquinarias del hogar, producción agrícola, pecuaria y forestal, así como los ingresos y gastos que derivan de éstos.

9. ORGANIZACIÓN DEL TRABAJO DE CAMPO

Las viviendas de la muestra son visitadas por Encuestadores, quienes recogen los datos referentes a cada una de las personas que habitan en las mismas, en un cuestionario destinado para cada hogar.

El trabajo de campo y el ingreso de datos son procesos descentralizados que se supervisan de manera muy estricta.

Cada equipo es dirigido por un supervisor y está compuesto por tres encuestadores, un digitador de datos y un chofer. Todos ellos están altamente capacitados para llevar a cabo el trabajo en terreno.

Duración de la entrevista: La entrevista tiene en promedio una duración de 2 horas, la cual se realiza en una o más visitas.

Informante de los datos: de preferencia la entrevista se realiza a cada miembro del hogar de 15 años y más de edad quien responde por sí mismo. En caso de menores de 15 años, podrán facilitar sus datos el Jefe o la Jefa del hogar o su cónyuge o algún miembro mejor informado sobre las actividades que aquellos realizan. En ningún caso los datos son suministrados por empleados domésticos, huésped, vecino, visitas, etc.

10. PROCESAMIENTO DE LOS DATOS

La digitación y edición de datos generalmente se llevan a cabo en terreno, en forma simultánea a la recolección, usualmente en las oficinas regionales.

Durante la digitación de datos, el programa de ingreso se encarga de realizar una serie de controles de calidad, revelando errores e inconsistencias para ser corregidos antes de abandonar el segmento.

Una vez asegurada la consistencia de toda la información en terreno, los datos se editan a nivel central para asegurar una vez más la calidad.

Se procede al cálculo de los factores de expansión, se confeccionan tablas y se presenta este resumen con los principales resultados de la encuesta, que contiene un conjunto cuidadosamente seleccionado de tablas.