

ASPECTOS METODOLÓGICOS DE LA ENCUESTA DE INGRESOS Y GASTOS Y DE CONDICIONES DE VIDA (EIGyCV) 2011-2012

1. OBJETIVO

El objetivo de la Encuesta de Ingresos y Gastos y de Condiciones de Vida 2011-2012 es actualizar la estructura de la Canasta Básica de Alimentos y la Canasta Total Familiar, cuyos valores constituyen las líneas de pobreza, así como el de caracterizar y analizar las Condiciones de Vida de la Población del Paraguay.

2. ÁMBITO GEOGRÁFICO

La Encuesta de Ingresos y Gastos 2011-2012 cubre todos los departamentos del país, excluyendo a los departamentos de Alto Paraguay y Boquerón, cuyas poblaciones representan menos del 2 % de la población total del país.

3. PERIODO DE LEVANTAMIENTO DE LOS DATOS

El levantamiento de los datos corresponde al período de agosto de 2011 a julio de 2012.

4. PERIODOS DE REFERENCIA

Los períodos de referencia para las variables a investigarse en la EIG y CV 2011/12, son los siguientes:

- ✓ Día de la entrevista.
- ✓ Últimos 7 días.
- ✓ Semana Pasada.
- ✓ 3 Semanas anteriores a la semana pasada.
- ✓ Últimos 30 días.
- ✓ Últimos 90 días.
- ✓ Últimos 3 meses.
- ✓ Últimos 12 meses.
- ✓ Año anterior.

5. 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.

6. MARCO MUESTRAL

El marco utilizado para la selección de la Muestra, es un marco de Unidades Primarias de Muestreo (**UPM**), llamadas "segmentos", los cuales fueron conformados a partir de los resultados del Censo Nacional de Población y Viviendas 2002. Los segmentos se definen a partir de una determinada cantidad de viviendas: 30 viviendas en promedio para el área urbana y 35 viviendas en promedio para el área rural, resultado de la unión de las áreas de empadronamiento del Censo 2002.

7. 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
- Los departamentos: San Pedro, Caaguazú, Itapúa, Alto Paraná y Central.

8. ESTRATIFICACIÓN GEOGRÁFICA

Las variables de estratificación son el departamento geográfico y el área de residencia. Para la EIG y CV se estableció un total de 13 estratos. Éstos comprenden Asunción, los 5 departamentos más grandes del país y un grupo que incluye el resto de los departamentos, divididos cada uno en estratos urbano y rural. Por lo tanto, la Encuesta está orientada a ser representativa de todo el país, por área urbano-rural y para los departamentos de San Pedro, Caaguazú, Itapúa, Alto Paraná y Central, más Asunción. Los departamentos de Boquerón y Alto Paraguay se excluyeron por razones de costos y tiempo.

9. TIPO DE MUESTREO

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

La muestra es probabilística porque las unidades han sido seleccionadas mediante métodos al azar, lo cual permite efectuar inferencias a la población en base a la teoría de probabilidades.

Es estratificada porque previamente a la selección, la población se ha dividido en estratos, con el objeto de mejorar su representatividad.

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).

9.1 SELECCIÓN DE VIVIENDAS

Primera etapa de selección: dentro de cada estrato, se selecciona un número determinado de UPM's en forma sistemática, con arranque aleatorio y con probabilidad proporcional al tamaño y la variabilidad de cada estrato, este método es conocido como afijación por Neyman. Más adelante se detallan los tamaños de muestra por estrato.

Segunda etapa de selección (Submuestreo): dentro de cada UPM escogida en la etapa anterior, se seleccionan 6 viviendas por (M.A.S) para el Área Urbana, y en el Área Rural se seleccionaron 12 viviendas distribuidos en 2 compactos de 6 viviendas cada una, iniciando con un arranque aleatorio dentro de cada UPM. En esta ocasión la selección fue sin reemplazo.

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

9.2 TAMAÑO Y DISTRIBUCIÓN DE LA MUESTRA

El tamaño de la muestra total previsto para la EIG y CV fue de 6.000 viviendas (definido por el presupuesto disponible). Para el Área Urbana se seleccionaron 658 UPM's y para el Área Rural fueron seleccionadas 171 UPM's y en cada una de ellas se investigaron 6 y 12 viviendas respectivamente.

En una propuesta inicial de muestreo se consideró una muestra total de 6.000 viviendas distribuidas en 500 UPM's, cada una constituida por 30 viviendas en promedio, de las cuales se seleccionan 12 viviendas, es decir más de un tercio de cada UPM. Se evaluó el coeficiente de variación (error de muestreo) del gasto mensual per cápita observado en la EIH 97/98, y al ser éste elevado en todos los estratos urbanos, se optó por aumentar el número de UPM's y reducir el número de viviendas seleccionadas al interior de ellas. Esto permitió reducir la varianza de los estimadores en dichos estratos.

Finalmente, se determinó el tamaño óptimo de los estratos considerando como variable de diseño el gasto mensual per cápita.

La muestra fue distribuida en estratos de la siguiente forma:

DISTRIBUCION DE LA MUESTRA POR ESTRATOS, EIGyCV (*)

ESTRATO	UPM URBANA	UPM RURAL	Total UPM	Total Viviendas
Asunción	144	-	144	864
San Pedro	22	31	53	504
Caaguazú	70	32	102	804
Itapúa	36	24	60	504
Alto Paraná	120	18	138	936
Central	175	15	190	1.230
Resto	91	51	142	1.158
Total	658	171	829	6.000

(*) Se excluyen los Departamentos de Boquerón y Alto Paraguay

10. CALIDAD DE LAS ESTIMACIONES

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

- **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.
- **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.).

11. 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 Stata utilizado actualmente se obtiene las estimaciones del error de muestreo, error relativo o coeficiente de variación, intervalos de confianza, de las variables más importantes de la Encuesta.

11.1 Estimación

Para valores absolutos, son datos muestrales multiplicados por el factor de expansión, para proporción o tasa son valores porcentuales.

11.2 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.

11.3 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.

11.4 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 z 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.

12. CUESTIONARIOS

CUESTIONARIO 1: consta de 9 secciones que tratan temas de educación, salud, empleo e ingreso de los miembros del hogar, miembros de la familia, gastos fuera del hogar:

CUESTIONARIO 2: capta los distintos gastos realizados en el hogar, los bienes duraderos con que cuenta el hogar, y en la última sección, en concordancia con las nuevas tendencias mundiales en cuanto a medición de la pobreza, incorpora una sección denominada “Otras dimensiones de la Pobreza”.

CUESTIONARIO 3: se recogen datos sobre Actividades Independientes no Agropecuarias, identificando el tipo de actividad económica que desempeña el encuestado o la empresa que dirige.

CUESTIONARIO 4: describe el perfil productivo y de ingreso de los hogares en los cuales algún miembro del hogar realiza de forma independiente actividad agropecuaria, forestal y derivados”.