

Elementos estadísticos básicos en la planeación de las encuestas

Andrés Gutiérrez

Asesor Regional en Estadísticas Sociales - Comisión Económica para América Latina y el Caribe (CEPAL) -
andres.gutierrez@cepal.org

Tabla de contenidos I

Introducción

Universo, muestra y unidades

Periodicidad en el tiempo

Tipos de encuestas en el tiempo

Parámetros e Indicadores de Interés

Indicadores compuestos y cambio neto

Introducción

Introducción

El fortalecimiento continuo de las investigaciones sociales es un objetivo que los institutos nacionales de estadística procuran cumplir de forma sistemática. En las encuestas que recolectan información primaria de hogares y personas, es fundamental documentar adecuadamente la metodología empleada. Este apartado explora distintos métodos de recolección de datos y aspectos clave en la planeación de una encuesta de hogares.

Universo, muestra y unidades

¿Qué es una encuesta?

- ▶ El término *encuesta* se refiere a la recolección de información de una población finita compuesta por individuos a observar y medir.
- ▶ La información se obtiene a través de entrevistas presenciales, telefónicas o formularios electrónicos autoadministrados.

El conjunto de estas unidades de interés se denomina *población objetivo* o *universo*.

Ejemplo:

La **Encuesta Nacional de Empleo y Desempleo** de Ecuador define su población objetivo como todas las personas mayores de 10 años residentes en viviendas particulares (INEC 2018).

Unidades de análisis

Las *unidades de análisis* son los niveles de desagregación definidos para estructurar el diseño y presentar los resultados.

- ▶ En México, la **ENIGH** clasifica las viviendas en:
 - ▶ Urbano alto
 - ▶ Complemento urbano
 - ▶ Rural
- ▶ En Colombia, la **Gran Encuesta Integrada de Hogares** usa 13 grandes ciudades con sus respectivas áreas metropolitanas como unidades de análisis (DANE 2017).

¿Por qué usar unidades de muestreo?

- ▶ En la práctica, no existe una lista completamente actualizada de todos los hogares de un país.
- ▶ Por eso, se seleccionan *unidades de muestreo* en varias etapas, que permiten acceder a la población objetivo.

Proceso de selección en etapas

1. Selección de zonas geográficas (UPM)
2. Selección de secciones cartográficas (USM)
3. Selección de hogares
4. Aplicación del cuestionario a un *respondiente idóneo*
5. (En algunos casos) Selección aleatoria de un individuo dentro del hogar (*unidad de observación*)

Ejemplo: Brasil - PNAD

La **Pesquisa Nacional por Amostra de Domicilios (PNAD)** de Brasil usa un diseño muestral en tres etapas:

- ▶ **UPM:** Municipios
- ▶ **USM:** Sectores censales (definidos en el Censo Demográfico)
- ▶ **Unidades finales:** Viviendas (IBGE-BR 2014)

Dinamismo de la población objetivo

Duncan y Kalton (1987, pág. 105) señalan que la población de interés cambia constantemente:

- ▶ Nacimientos
- ▶ Defunciones
- ▶ Migración
- ▶ Institucionalización (cárceles, hospitales, etc.)

La población objetivo de las encuestas de hogares en América Latina suele excluir a personas en instituciones colectivas o militares.

Implicaciones para el diseño

Dado que los hogares se crean y desaparecen con rapidez, los equipos técnicos de las Oficinas Nacionales de Estadística deben:

- ▶ Reconocer que los objetivos de la encuesta pueden permanecer,
- ▶ Pero la **población objetivo cambia con el tiempo**.

Por ello, es clave definir mecanismos de seguimiento y actualización que reflejen esta dinámica social.

Periodicidad en el tiempo

Encuestas como herramienta continua

- ▶ Las Oficinas Nacionales de Estadística (ONE) no realizan encuestas de forma aislada.
- ▶ Estas operaciones son fundamentales para el seguimiento de indicadores clave y muchas se implementan de forma **periódica** o **continua**.
- ▶ Una buena planeación considera esta continuidad, lo cual permite:
 - ▶ Recolección más eficiente
 - ▶ Estimaciones ajustadas a los recursos disponibles
 - ▶ Mejores decisiones de política pública

¿Por qué no basta con una medición única?

Si la población fuera estática, una encuesta transversal sería suficiente.

Pero como la población cambia constantemente, se requiere una planificación dinámica.

Factores de cambio:

- ▶ Nacimientos, muertes, migraciones
- ▶ Cambios en el hogar (división, unión, disolución)
- ▶ Cambios en edad o condición de los individuos

Encuestas según su frecuencia

1. Encuestas esporádicas:

Se aplican en puntos específicos del tiempo, cuando una sola medición es suficiente.

Ejemplo:

Encuestas de Ingresos y Gastos

- ▶ Periodicidad: cada 5 años aprox.
- ▶ Objetivo: actualizar la canasta básica, medir pobreza (CEPAL 2018).

Encuestas según su frecuencia

2. Encuestas periódicas o continuas:

Se realizan con intervalos regulares para monitorear dinámicas sociales y económicas.

Ejemplo:

Encuestas de empleo o mercado laboral

- ▶ Cambios frecuentes requieren medición constante
- ▶ Una única medición no basta para el análisis de tendencias

Retos y beneficios de encuestas continuas

Retos:

- ▶ Diseños más complejos
- ▶ Necesidad de mantener marcos actualizados
- ▶ Mayor demanda logística y presupuestaria

Beneficios:

- ▶ Información más oportuna
- ▶ Mejora en la calidad del seguimiento estadístico
- ▶ Mayor capacidad para responder a cambios sociales

El tiempo y la población

La estructura poblacional cambia con el tiempo. Por ejemplo:

- ▶ Nuevos hogares: jóvenes que se independizan
- ▶ Hogares que desaparecen: fallecimientos, migraciones
- ▶ Hogares que se transforman: divorcios, uniones

Por eso, las unidades de observación deben ser incorporadas o excluidas según cambios naturales en la población.

Recomendación

- ▶ Tener en cuenta la **periodicidad en el tiempo** es esencial para diseñar encuestas útiles y representativas.
- ▶ La población cambia, y el sistema estadístico debe adaptarse a ese dinamismo.

Tipos de encuestas en el tiempo

Encuestas transversales

- ▶ Este tipo de encuestas recolecta información en un único punto específico en el tiempo o sobre un periodo de referencia limitado.
- ▶ La información recolectada se refiere exclusivamente a ese momento, por lo tanto:
 - ▶ No permite estimar cambios en el tiempo
 - ▶ No es posible hacer comparaciones intertemporales (excepto mediante preguntas retrospectivas)
 - ▶ La muestra no se repite en otros periodos

Ejemplo: Encuesta transversal

La siguiente tabla ilustra cómo una muestra específica de hogares es observada únicamente en el **Tiempo 2**:

Tabla 1: *Esquema de una encuesta transversal.*

Hogar	Tiempo 1	Tiempo 2	Tiempo 3	Tiempo 4	...	Tiempo T
1		x				
2		x				
3		x				
4		x				
...		x				
n		x				

Encuestas repetidas

- ▶ Cuando el interés es **dar seguimiento en el tiempo** a una población, se usan encuestas repetidas.
- ▶ Estas permiten observar la evolución del fenómeno bajo estudio, pero **no permiten analizar trayectorias individuales** porque:
 - ▶ Cada muestra es distinta
 - ▶ No se hace seguimiento a los mismos hogares

Ejemplo: Encuesta repetida

En cada levantamiento se selecciona una muestra distinta en cada tiempo:

Tabla 2: *Esquema de una encuesta repetida*

Hogar	Tiempo 1	Tiempo 2	Tiempo 3	Tiempo 4	...	Tiempo T
1	x	x	x	x	x	x
2	x	x	x	x	x	x
3	x	x	x	x	x	x
4						
...						
n						

Encuestas de panel dividido

Para hacer frente a las limitaciones de las encuestas de panel y capturar tanto los **cambios individuales** como los **cambios estructurales de la población**, se diseñan las **encuestas de panel dividido**.

Este tipo de levantamiento combina:

- ▶ **Panel puro** (seguimiento de los mismos hogares)
- ▶ **Muestra repetida** (incorporación de nuevos hogares en cada periodo)

Encuestas de panel dividido

Para hacer frente a las limitaciones de las encuestas de panel y capturar tanto los **cambios individuales** como los **cambios estructurales de la población**, se diseñan las **encuestas de panel dividido**.

Este tipo de levantamiento combina:

- ▶ **Panel puro** (seguimiento de los mismos hogares)
- ▶ **Muestra repetida** (incorporación de nuevos hogares en cada periodo)

Este enfoque permite:

- ▶ Estudiar trayectorias individuales a lo largo del tiempo
- ▶ Analizar la evolución de la población actual
- ▶ Mejorar la precisión de las inferencias para cohortes y estructura poblacional

Ejemplo: Encuesta de panel dividido

Tabla 3: *Esquema de una encuesta de panel dividido*

Hogar	Tiempo 1	Tiempo 2	Tiempo 3	Tiempo 4	...	Tiempo T
1	x	x	x	x	x	x
2	x					
3		x				
4			x			
5				x		
...					x	
n						x

Encuestas de panel rotativo

Las encuestas de panel rotativo se diseñan para balancear el seguimiento individual y la captación de cambios estructurales en la población a lo largo del tiempo.

Estas encuestas:

- ▶ Mantienen parcialmente hogares de periodos anteriores (traslape parcial)
- ▶ Incorporan nuevas unidades de observación
- ▶ Reducen el desgaste del panel y la pérdida de muestra
- ▶ Permiten inferencias dinámicas sin incurrir en el costo completo de un panel puro

Encuestas de panel rotativo

Las encuestas de panel rotativo se diseñan para balancear el seguimiento individual y la captación de cambios estructurales en la población a lo largo del tiempo.

Estas encuestas:

- ▶ Mantienen parcialmente hogares de periodos anteriores (traslape parcial)
- ▶ Incorporan nuevas unidades de observación
- ▶ Reducen el desgaste del panel y la pérdida de muestra
- ▶ Permiten inferencias dinámicas sin incurrir en el costo completo de un panel puro

A diferencia de un panel clásico (con traslape total), aquí el solapamiento se da solo entre algunos periodos consecutivos, reduciendo la fatiga del respondente y manteniendo la frescura de la muestra.

Ejemplo: Diseño de panel rotativo

Tabla 4: *Esquema de una encuesta de panel rotativo.*

Hogar	Tiempo 1	Tiempo 2	Tiempo 3	Tiempo 4	Tiempo 5	Tiempo 6
1	x					
2	x	x				
3	x	x	x			
4		x	x	x		
5			x	x	x	
6				x	x	x
...					x	x
n						x

Rotación de paneles

Algunas encuestas de hogares en América Latina permiten que los hogares sean entrevistados en más de una ocasión, para captar cambios en su situación socioeconómica. Este seguimiento es fundamental para estimar cambios de estado (como empleo, pobreza, etc.), lo cual solo es posible a través de diseños con componente de panel.

Propiedades del diseño rotacional $2(2)_2$

Por ejemplo, la *Encuesta Permanente de Hogares* en Argentina emplea un diseño de rotación $2(2)_2$, en el cual las viviendas:

- ▶ Son entrevistadas durante dos trimestres consecutivos
- ▶ Excluidas durante los dos siguientes
- ▶ Reincorporadas por otros dos trimestres

Este esquema cubre un período de seguimiento de hasta 18 meses, permitiendo la estimación eficiente de indicadores dinámicos.

Propiedades del diseño rotacional $2(2)^2$

- ▶ Traslape del 50% entre trimestres consecutivos (ej. T1 y T2)
- ▶ Cero traslape entre trimestres alternos (ej. T1 y T3)
- ▶ Traslape del 25% en trimestres más alejados (ej. T1 y T4)
- ▶ Traslape del 50% al cabo de un año (ej. T1 2016 y T1 2017)

Propiedades del diseño rotacional $2(2)2$

Tabla 5: *Rotación de paneles en un diseño $2(2)2$.*

Año	Trimestre	A	B	C	D
2016	T1	$a1$	$b1$	$c1$	$d1$
	T2	$a1$	$b2$	$c2$	$d1$
	T3	$a2$	$b2$	$c2$	$d2$
	T4	$a2$	$b1$	$c3$	$d2$
2017	T1	$a1$	$b1$	$c3$	$d3$
	T2	$a1$	$b2$	$c4$	$d3$
	T3	$a2$	$b2$	$c4$	$d4$
	T4	$a2$	$b3$	$c3$	$d4$
2018	T1	$a3$	$b3$	$c3$	$d3$
	T2	$a3$	$b4$	$c4$	$d3$
	T3	$a4$	$b4$	$c4$	$d4$
	T4	$a4$	$b3$	$c5$	$d4$

Encuesta Continua de Empleo: Diseño 4(0)1

En Bolivia, el Instituto Nacional de Estadística aplica la *Encuesta Continua de Empleo* bajo un esquema rotacional 4(0)1 en áreas urbanas. En este diseño:

- ▶ Una vivienda es entrevistada durante 4 trimestres consecutivos
- ▶ Luego, se excluye del panel permanentemente

Este enfoque permite estimar dinámicas del mercado laboral urbano, como la tasa de desempleo y subocupación, manteniendo representatividad temporal.

Propiedades del diseño 4(0)1

- ▶ Traslape del 75% entre trimestres consecutivos (ej. T1 y T2)
- ▶ Traslape del 50% entre trimestres alternos (ej. T1 y T3)
- ▶ Traslape del 25% en trimestres más lejanos (ej. T1 y T4)
- ▶ Sin traslape a partir del quinto periodo (ej. T1 y T5)

Propiedades del diseño 4(0)1

Tabla 6: *Rotación de paneles en un diseño 4(0)1 aplicado por el INE-Bolivia.*

Año	Trimestre	A	B	C	D
2016	T1	<i>a1</i>	<i>b1</i>	<i>c1</i>	<i>d1</i>
	T2	<i>a1</i>	<i>b2</i>	<i>c1</i>	<i>d1</i>
	T3	<i>a1</i>	<i>b2</i>	<i>c2</i>	<i>d1</i>
	T4	<i>a1</i>	<i>b2</i>	<i>c2</i>	<i>d2</i>
2017	T1	<i>a2</i>	<i>b2</i>	<i>c2</i>	<i>d2</i>
	T2	<i>a2</i>	<i>b3</i>	<i>c2</i>	<i>d2</i>
	T3	<i>a2</i>	<i>b3</i>	<i>c3</i>	<i>d2</i>
	T4	<i>a2</i>	<i>b3</i>	<i>c3</i>	<i>d3</i>
2018	T1	<i>a3</i>	<i>b3</i>	<i>c3</i>	<i>d3</i>
	T2	<i>a3</i>	<i>b4</i>	<i>c3</i>	<i>d3</i>
	T3	<i>a3</i>	<i>b4</i>	<i>c4</i>	<i>d3</i>
	T4	<i>a3</i>	<i>b4</i>	<i>c4</i>	<i>d4</i>

Parámetros e Indicadores de Interés

Tipos de Indicadores (**Indicadores descriptivos:**)

Las encuestas permiten estimar parámetros que describen la situación de una población. Estos se agrupan en dos grandes tipos:

- ▶ **Medias:** ingreso per cápita, gasto mensual, años de escolaridad
- ▶ **Proporciones:** Porcentaje de personas en pobreza extrema, Porcentaje de viviendas con pisos de tierra, Porcentaje de niños con desnutrición
- ▶ **Totales:** gasto total en alimentos, total de remesas
- ▶ **Tamaños poblacionales:** cantidad de hogares en condición de pobreza, personas desempleadas, mujeres víctimas de violencia laboral

Tipos de Indicadores (**Indicadores analíticos:**)

- ▶ **Correlaciones:** relación entre lectura y escolaridad
- ▶ **Regresiones:** cambio esperado en ingreso según experiencia laboral

Total poblacional y tamaño de la población

Gran parte del análisis estadístico se basa en estimar **totales poblacionales**, definidos como:

$$t_y = \sum_{k \in U} y_k$$

Donde:

- ▶ U : conjunto de unidades de la población
- ▶ y_k : variable de interés para la unidad k

Ejemplo: Estimar el gasto total.

- ▶ Estimar el gasto total de los hogares en alimentos y bebidas no alcohólicas.
 - ▶ U : todos los hogares del país
 - ▶ y_k : gasto del hogar k en el rubro alimentario

Ejemplo: Estimar el gasto total.

- ▶ Estimar el gasto total de los hogares en alimentos y bebidas no alcohólicas.
 - ▶ U : todos los hogares del país
 - ▶ y_k : gasto del hogar k en el rubro alimentario
- ▶ Un caso especial es el **tamaño poblacional**:

$$N = \sum_{k \in U} 1$$

- ▶ N : cantidad total de unidades en la población

Usualmente aproximado mediante censos o proyecciones oficiales

Promedio poblacional

Uno de los parámetros más utilizados en investigación social es el **promedio poblacional**, que representa una asignación equitativa de la variable de interés:

$$\bar{y}_U = \frac{t_y}{N}$$

Donde:

- ▶ t_y : total poblacional de la variable de interés
- ▶ N : tamaño de la población

Por ejemplos, el ingreso promedio por hogar y Gasto promedio en educación

El promedio permite comparaciones entre grupos, regiones o periodos de tiempo.

Parámetros sobre dominios poblacionales

En muchos estudios interesa analizar subpoblaciones o **dominios** $U_d \subseteq U$. Los parámetros clave son:

Tamaño del dominio:

Cantidad de unidades con una característica específica:

$$N_d = \sum_{k \in U} z_{d_k} = \sum_{k \in U_d} 1$$

Ejemplo: número de personas desempleadas en el país.

Parámetros sobre dominios poblacionales (**Proporción del dominio**)

Proporción del total de la población que pertenece al dominio:

$$P_d = \frac{N_d}{N}$$

Ejemplo: Porcentaje de personas en pobreza extrema.

Parámetros sobre dominios poblacionales (**Total del dominio:**)

Suma de la variable de interés dentro del dominio:

$$t_{y_d} = \sum_{k \in U} y_k \cdot z_{d_k} = \sum_{k \in U_d} y_k$$

Ejemplo: ingreso total de las mujeres, gasto total en educación en hogares rurales.

Media del dominio

Cantidad promedio de la variable y dentro de un dominio U_d :

$$\bar{y}_{U_d} = \frac{t_{y_d}}{N_d}$$

Ejemplos:

- ▶ Ingreso promedio entre mujeres
- ▶ Gasto promedio en comida en la población indígena

Razón poblacional:

Relación entre dos totales poblacionales:

$$R_U = \frac{t_y}{t_x}$$

Ejemplos:

- ▶ Tasa de desocupación = $\frac{\text{Total de desocupados}}{\text{Total de activos}}$
- ▶ Ingreso per cápita = $\frac{\text{Total ingresos}}{\text{Total de personas}}$

Estas razones permiten construir múltiples indicadores sociales.

Indicadores compuestos y cambio neto

Índices de pobreza (Foster, Greer y Thorbecke, 1984):

$$F_{\alpha} = \frac{1}{N} \sum_U \left(\frac{u - y_k}{u} \right)^{\alpha} I_{(y_k < u)}$$

Donde:

- ▶ y_k : ingreso del individuo k
- ▶ u : umbral de pobreza
- ▶ α : sensibilidad del índice

Casos particulares:

- ▶ $\alpha = 0$: Incidencia (tasa de pobreza)
- ▶ $\alpha = 1$: Brecha de pobreza
- ▶ $\alpha = 2$: Severidad de pobreza

Cambio neto entre periodos:

Diferencia entre totales en dos momentos del tiempo:

$$\Delta_y = t_{y(2)} - t_{y(1)}$$

En donde $t_{y(2)}$ es el total de interés en el tiempo $t = 2$, y $t_{y(1)}$ lo es en el tiempo $t = 1$.

Ejemplo: cambio en el gasto en salud entre 2023 y 2024

Cambios netos en el mercado laboral

Ejemplo (cifras en miles):

► Empleados: $9699 - 10012 = -313$

► Desempleados: $614 - 694 = -80$

► Inactivos: $6754 - 6361 = +393$

Tabla 7: *Composición del mercado de trabajo en dos periodos de tiempo (cifras en miles de personas). Las columnas corresponden al segundo periodo y las filas al primero.*

Condición	Ocupado	Desocupado	Inactivo	Total
Ocupado	9222	128	662	10012
Desocupado	221	322	151	694
Inactivo	256	164	5941	6361
Total	9699	614	6754	17067

Cambios brutos en el mercado laboral

Ejemplo, flujos entre estados (tabla cruzada):

- ▶ 92.1% de empleados conservaron su empleo
- ▶ 4.0% de inactivos pasaron a estar ocupados
- ▶ 1.3% de ocupados perdieron su empleo
- ▶ 46.4% de desempleados siguieron desempleados

Tabla 8: *Composición del mercado de trabajo en dos periodos de tiempo (cifras en miles de personas). Las columnas corresponden al segundo periodo y las filas al primero.*

Condición	Ocupado	Desocupado	Inactivo	Total
Ocupado	9222	128	662	10012
Desocupado	221	322	151	694
Inactivo	256	164	5941	6361
Total	9699	614	6754	17067

¡Gracias!

Email: andres.gutierrez@cepal.org

Referencias

- CEPAL. 2018. *Medición de la pobreza por ingresos - Actualización metodológica y resultados*. Metodologías de la CEPAL.
http://repositorio.cepal.org/bitstream/handle/11362/44314/1/S1800852_es.pdf.
- DANE. 2017. «Gran Encuesta Integrada de Hogares - - Departamento Administrativo Nacional de Estadística».
http://formularios.dane.gov.co/Anda_4_1/index.php/catalog/458.
- Duncan, G. J., y G. Kalton. 1987. «Issues of Design and Analysis of Surveys across Time». *International Statistical Review / Revue Internationale de Statistique* 55 (1): 97. <https://doi.org/10.2307/1403273>.
- IBGE-BR. 2014. «Pesquisa Nacional por Amostra de Domicílios Continua - Notas Metodológicas».
- INEC. 2018. «Instituto Nacional de Estadística y Censos».
<http://www.ilo.org/surveydata/index.php/catalog/1393/study-description>.