

DISEÑO DE MUESTRA ENCUESTA DE HOGARES 2018

UNIDAD DE MUESTREO Y LOGISTICA

Contenido

1. Introducción	1
2. Aspectos Metodológicos de la Encuesta de Hogares 2018.....	1
2.1. Objetivo General	1
2.2. Objetivos Específicos	1
3. Diseño Muestral	2
3.1 Universo de Estudio.....	2
3.2 Ámbito Geográfico	2
3.3 Unidades de Observación y Análisis	2
3.4 Unidades de Muestreo.	2
3.5 Periodo de levantamiento de los datos.....	3
4. Marco Muestral	3
4.1 Criterios de Conglomeración y Estratificación del Marco Muestral	4
Variable de Conglomeración.....	4
Variable de Estratificación	4
Variable de Estratificación Estadística Socio-Económico	5
5. Características generales del diseño de muestra.....	6
6. Nivel de desagregación de la información.....	6
7. Cálculo del Tamaño de Muestra.....	6
8. Distribución de la Muestra por Departamento y área geográfica	7
9. Factores de Expansión	8
10. Ajustes a los Factores Expansión	8
11. Estimadores y Errores de Muestreo.....	9
12. Incidencia de campo	9

Contenido de tablas

Tabla 1: Estructura del Marco Muestral	3
Tabla 2: Marco Muestral 2012	4
Tabla 3: Marco Muestral 2012 -Estrato Geográfico	5
Tabla 4: Reagrupación geográfica	5
Tabla 5: Estrato Estadístico.....	5
Tabla 6: Coeficiente de variación (CV) del tamaño de muestra planificada	7
Tabla 7: Distribución de la muestra planificada de viviendas y UPM	7
Tabla 8: Coeficiente de variación (CV) del tamaño de muestra ejecutada.....	10
Tabla 9: Incidencia de campo por departamento.....	11

1. Introducción

Las encuestas de hogares se desarrollaron a partir de 1980 con las encuestas permanentes de hogares, las cuales se remplazaron por las encuestas integradas de hogares EIH, a partir de 1999 a 2002 se desarrollaron encuestas puntuales del programa para el mejoramiento de las encuestas y la Medición de Condiciones de Vida MECOVI. Con la culminación del programa MECOVI, se inició la encuesta continua de hogares durante 2003 y 2004, incluyendo temas generales del programa MECOVI así como el de presupuestos familiares y un módulo sobre gestión de gobierno y democracia, propuesto en las reuniones de trabajo de la comunidad Andina de Naciones. Desde el 2005 a 2009 se retoma las encuestas puntuales de corte transversal, manteniendo como temática de exploración "las condiciones de vida de la población" donde los resultados se muestran a nivel nacional, total urbano y total rural.

A partir del 2011 se realiza la desagregación a nivel Nacional, Nacional Urbano, Nacional Rural y departamental, donde los tamaños de muestra son calculados de manera independiente por departamento excepto para los departamentos de Beni y Pando, donde las estimaciones deben considerarse de manera conjunta.

Es en la gestión 2016, en cumplimiento de la Ley Nro. 548 y la política de fortalecimiento y actualización de las estadísticas e indicadores sobre trabajo de Niñas, Niños y Adolescentes, el Instituto Nacional de Estadística, adiciona a la encuesta de hogares la Encuesta de Niñas, Niños y Adolescentes que realizan una Actividad Laboral o Trabajan "EH-ENNA 2016" en coordinación con el Ministerio de Trabajo, Empleo y Prevención Social y el Ministerio de Planificación y Desarrollo.

Es en la gestión 2016, en cumplimiento de la Ley Nro. 548 y la política de fortalecimiento y actualización de las estadísticas e indicadores sobre trabajo de Niñas, Niños y Adolescentes, el Instituto Nacional de Estadística, adiciona a la encuesta de hogares la Encuesta de Niñas, Niños y Adolescentes que realizan una Actividad Laboral o Trabajan "EH-ENNA 2016" en coordinación con el Ministerio de Trabajo, Empleo y Prevención Social y el Ministerio de Planificación y Desarrollo.

En la gestión 2017 y 2018 para la "EH 2017" y "EH 2018" se incluyó el módulo de Culturas, y para la EH 2018 se incluyeron preguntas referidas a discriminación y cambios para la medición de personas con discapacidad

2. Aspectos Metodológicos de la Encuesta de Hogares 2018

2.1. Objetivo General

- Suministrar información sobre las condiciones de vida de los hogares, a partir de la recopilación de información de variables socioeconómicas y demográficas de la población boliviana, necesaria para la formulación, evaluación, seguimiento de políticas y diseño de programas de acción en el área social.

2.2. Objetivos Específicos

La Encuesta de Hogares 2018 tiene los siguientes objetivos específicos:

- Generar indicadores relacionados con empleo, desempleo y otras características económicas, que permitirán conocer la evolución de la pobreza, el bienestar y la tendencia de las

condiciones de vida de los hogares del país, con resultados a nivel nacional, con desagregación nacional urbano / rural y a nivel departamental, excepto Beni y Pando que será de manera conjunta.

- Generar indicadores de ocupación para las ciudades capitales de los departamentos del eje central (La Paz, Cochabamba y Santa Cruz).
- Servirá de fuente de información a instituciones públicas y privadas que permita realizar investigaciones futuras.

3. Diseño Muestral

Con el fin de lograr la comparabilidad, el diseño muestral de la Encuesta de Hogares 2018 mantiene similitud a los diseños empleados en las encuestas precedentes de las EH.

3.1 Universo de Estudio

La investigación está dirigida al conjunto de hogares establecidos en viviendas particulares ocupadas dentro de los nueve (9) departamentos del país.

Se excluye de la investigación a la población residente en viviendas colectivas tales como hospitales, hoteles, cárceles, recintos militares, hogares de ancianos, unidades educativas, etc.; se incluyen a hogares que, forman un grupo independiente y además residen dentro de estos establecimientos, ejemplo: porteros, cuidadores, etc.

3.2 Ámbito Geográfico

El ámbito geográfico de la encuesta de hogares 2018, abarca el área urbana y rural de todos los departamentos del país: Chuquisaca, La Paz, Cochabamba, Oruro, Potosí, Tarija, Santa Cruz, Beni y Pando.

3.3 Unidades de Observación y Análisis

Unidades de observación.- La investigación está dirigida al conjunto de hogares establecidos en viviendas particulares ocupadas dentro del territorio de Bolivia en la gestión 2018.

Unidades de análisis.- Estos se constituirán en los hogares, y miembros del hogar para cuales se releva información.

Se excluye de la investigación a la población residente en viviendas colectivas tales como hospitales, hoteles, cárceles, recintos militares, hogares de ancianos, unidades educativas, etc.; se incluyen a hogares que, forman un grupo independiente y además residen dentro de estos establecimientos, ejemplo: porteros, cuidadores, etc.

3.4 Unidades de Muestreo.

La unidad de muestreo, en su última etapa, es la vivienda particular ocupada, ésta tiene permanencia fija en el tiempo y espacio, característica que la habilita para ser utilizada como unidad de selección en

el diseño muestral. La unidad de muestreo, de acuerdo a las áreas de estudio, se conforma de la siguiente manera:

Las unidades de muestreo son:

- UPM - La unidad primaria de muestreo es el conglomerado de viviendas.
- USM - La unidad secundaria de muestreo conforman las viviendas particulares dentro de las UPM seleccionadas.

3.5 Periodo de levantamiento de los datos

El levantamiento de información primaria corresponde al periodo del 22 de octubre al 21 de diciembre de 2018.

4. Marco Muestral

La información del Marco Muestral (MM) está basada en el Censo de Población y Vivienda de 2012 (CNPV-2012), la Actualización Cartográfica Multipropósito (ACM-2010-2012) y el Censo Nacional Agropecuario de 2013 (CNA-2013).

El MM-2012 cubre todo el territorio nacional incluyendo las áreas de difícil acceso. Codifica las áreas geográficas en unidades primarias de muestreo (UPM), constituyéndose bien delimitadas en el área rural dispersa; área rural amanzanada; área urbana y periferias (centros poblados que se encuentran en el área urbana y rural a la vez). Las unidades secundarias de muestreo (USM)

Están designa a las viviendas, cubriendo la totalidad de las viviendas particulares¹.

La conformación de las Unidades Primarias de Muestreo (UPM), en el área urbana y rural, está realizada de acuerdo con los límites geográficos y estadísticos bajo los siguientes criterios de Agrupación y División

La estructura del Marco Muestral está definida por las siguientes variables:

Tabla 1: Estructura del Marco Muestral

Tipo de Variable	Nombre de la variable	Variable
Variables de División Política Administrativa	I02_DEPTO	Departamento: código y descripción
	I03_PROV	Provincia: código y descripción
	NOMBRE_PROVINCIA	
	I04_MUN	Municipio: código y descripción
	NOMBRE_MUNICIPIO	
Variables de organización municipal administrativa rural	I04_MUN	Municipio: código y descripción
	NOMBRE_MUNICIPIO	
	I06_CIUCOM	Comunidad: código y descripción
	NOMBRE_COMUNIDAD	

¹ Vivienda particular es aquella que está habitada hasta por tres hogares, una vivienda con más de tres hogares es reconocida como vivienda colectiva (por ejemplo: hospitales, cuarteles, etc.)

Variables de Organización Municipal administrativa urbana	I04_MUN	Municipio: código y descripción
	NOMBRE_MUNICIPIO	
	ID_UNICO	
Variables para el diseño de muestra	ID_UNICO_UPM	UPM dado por la unidad de cartografía
	upm	UPM normalizado
	NUM_UPM	UPM dado por la unidad de Muestreo
	Categoría_URBRUR	Estrato geográfico
	URBRUR	Área Urbana o Rural
	Vivienda_ocu	Número de viviendas Ocupadas
	totperso	Total de persona
	totalHombre	Total de personas hombres
	totalMujer	Total de personas mujeres

Fuente: Instituto Nacional de Estadística

Los planos y mapas cartográficos constituyen parte del marco muestral para la encuesta, mediante el cual se identifican las áreas correspondientes a las unidades primarias de muestreo seleccionadas y sirven de guía para ubicación en el terreno y el desarrollo del plan de recorridos de brigadas de encuestadores y supervisores.

4.1 Criterios de Conglomeración y Estratificación del Marco Muestral

Variable de Conglomeración

A partir del 2014 para las encuestas de hogares se utilizó el Marco Muestral 2012 que está basado en la información del Censo de Población y Vivienda CNPV-2012 y está constituido por 102'593 conglomerados (UPM) a nivel nacional y distribuidos por departamento (ver tabla 2).

Tabla 2: Marco Muestral 2012
Distribución de los conglomerados o UPM por departamento

Departamento	Casos	Porcentaje
Total Bolivia	102'593	100,0%
Chuquisaca	5'362	5,2%
La Paz	29'610	28,9%
Cochabamba	18'041	17,6%
Oruro	7'384	7,2%
Potosí	8'519	8,3%
Tarija	4'578	4,5%
Santa Cruz	23'690	23,1%
Beni	4'350	4,2%
Pando	1'059	1,0%

Fuente: Instituto Nacional de Estadística

Variable de Estratificación

Con el propósito de obtener precisión en los estimadores en el sistema de encuestas a Hogares se realizó la estratificación en base a características socioeconómicas de manera que los conglomerados sean homogéneos dentro de los grupos y heterogéneos entre ellos.

El estrato geográfico está definido por las siguientes características:

Tabla 3: Marco Muestral 2012 -Estrato Geográfico

Estrato Geográfico	Descripción
1	Ciudades Capitales
2	Ciudades Intermedias (10'000 y más habitantes)
3	Centros Poblados (2'000 a 9'999 habitantes)
4	Localidades con características ameznadas (250 a 1'999 habitantes)
5	Áreas Dispersas (menos de 250 habitantes)
6	Ciudad de El Alto
7	Áreas Dispersas (mayor a 2'000 habitantes)

Fuente: Instituto Nacional de Estadística

El estrato geográfico por área urbana y rural está definido:

Tabla 4: Reagrupación geográfica

Estrato Geográfico	Descripción	Reagrupación geográfica
1	Ciudades Capitales	1. Urbano
2	Ciudades Intermedias (10'000 y más habitantes)	
3	Centros Poblados (2'000 a 9'999 habitantes)	
6	Ciudad de El Alto	
4	Localidades con características ameznadas (250 a 1'999 habitantes)	2. Rural
5	Áreas Dispersas (menos de 250 habitantes)	
7	Áreas Dispersas (mayor a 2'000 habitantes)	

Fuente: Instituto Nacional de Estadística

Variable de Estratificación Estadística Socio-Económico

Para la construcción de la variable de estratificación estadística socio-económica se consideró variables estructurales de la base del Censo de Población y Vivienda 2012, entre ellas: Características de la vivienda; Acceso a Servicios Básicos; Hacinamiento; Capital Humano; Educación; Patrimonio de los hogares; y Capacidad para la generación de ingresos.

Para la encuesta EH-2018 se considera la combinación de estas variables (Estrato Geográfico y Estrato Estadístico (socio-económico)), que permite corregir las variaciones de las estimaciones de los estratos.

Tabla 5: Estrato Estadístico

Estrato Estadístico	Descripción
1	Bajo
2	Medio Bajo
3	Medio Alto
4	Alto

Fuente: Instituto Nacional de Estadística

5. Características generales del diseño de muestra.

El diseño muestral aplicado para la EH 2018, tiene las características de un muestreo probabilístico, estratificado por conglomerados y bi etapico.

Probabilístico, porque la probabilidad de selección es conocida y diferente de cero para cada elemento objeto de estudio, lo cual permite realizar inferencia estadística al total de la población objetivo.

Por conglomerados, debido a que las unidades objeto de estudio (viviendas/hogares) se encuentran agrupadas en áreas conformadas, con fines de muestreo, en base a la información y cartografía del Censo de Población y Vivienda 2012, denominadas conglomerados.

Estratificado, porque previo a la selección de conglomerados, los estratos son clasificados con el propósito de mejorar la precisión de los estimadores y permitir un mejor control de la distribución de la muestra.

Bi-etapico, porque los elementos pertenecientes a la muestra se seleccionan en dos etapas:

- Etapa 1: Selección de conglomerados UPM (manzanas o comunidades)
- Etapa 2: Selección de viviendas USM, al interior de la UPM seleccionada o elegida.

6. Nivel de desagregación de la información

Se garantiza estimaciones de indicadores de pobreza, con errores muestrales admisibles que permitan conocer la evolución de estas variables, el nivel de desagregación es:

- Nacional
- Nacional Urbano
- Nacional Rural
- Departamental (En el caso de Beni y Pando se obtendrán estimaciones conjuntas)

Presentar indicadores de ocupación para las siguientes ciudades capitales: La Paz, Cochabamba y Santa Cruz.

Realizar desagregaciones mayores a los planteados en el diseño, estará sujeto a la valoración de los errores muestrales.

7. Cálculo del Tamaño de Muestra

Las variables que han sido, y son, determinantes del cálculo de tamaños muestrales apropiados para cada estrato y dominio de estudio son: la **tasa de desempleo y pobreza**. Los datos de referencia para el cálculo del tamaño de muestra fueron en base a indicadores de pobreza y empleo obtenidos con la información de la Encuesta de Hogares 2017.

Para garantizar la precisión en los estimadores, se procedió de manera independiente en cada uno de los dominios de estudio. En el caso de las ciudades capitales de los departamentos de La Paz, Cochabamba y Santa Cruz se empleó el coeficiente de variación de la tasa de desempleo; en el resto se consideró el indicador de pobreza.

Tabla 6: Coeficiente de variación (CV) del tamaño de muestra planificada

Departamento/Ciudad		Número de viviendas (muestra planificada 2018)	Coeficiente de Variación esperado (CV) en base a la Gestión 2017		
			CV de desempleo	CV de pobreza	CV de pobreza extrema
Nacional		11'448	2,7	1,4	2
Chuquisaca	Chuquisaca	624		4,4	5,6
La Paz	La Paz (Ciudad)	1104	6,9	6,3	11,1
	La Paz (Resto)	852		3,9	5,3
	La Paz (El Alto)	1104		5	7,7
Cochabamba	Cochabamba (Ciudad)	1452	6	4	6,6
	Cochabamba (Resto)	840		4,4	5,8
Oruro	Oruro	696		5,7	8,6
Potosí	Potosí	600		4,6	5,6
Tarija	Tarija	1020		4,7	7,4
Santa Cruz	Santa Cruz (Ciudad)	1236	7,9	5,9	11,8
	Santa Cruz (Resto)	804		5,8	8,9
Beni	Beni	660		5,2	7,3
Pando	Pando	456		7,5	10,7

Fuente: Instituto Nacional de Estadística

8. Distribución de la Muestra por Departamento y área geográfica

La Tabla 7, presenta la distribución de la muestra de viviendas por departamento y área geográfica, así como la distribución de la UPM, en cada una de las UPM se selecciona 12 viviendas.

**Tabla 7: Distribución de la muestra planificada de viviendas y UPM
Por departamento y área geográfica**

	Número de Viviendas de la Muestra				Numero de UPM de la Muestra			
	Total	Urbano	Resto Urbano	Rural	Total	Urbano	Resto Urbano	Rural
NACIONAL	11.448	6.396	2472	2580	954	533	206	215
Chuquisaca	624	276	72	276	52	23	6	23
La Paz	3.060	2.208	444	408	255	184	37	34
Cochabamba	2.292	1.452	468	372	191	121	39	31
Oruro	696	324	84	288	58	27	7	24
Potosí	600	120	156	324	50	10	13	27
Tarija	1.020	408	324	288	85	34	27	24
Santa Cruz	2.040	1.236	540	264	170	103	45	22
Beni	660	144	348	168	55	12	29	14
Pando	456	228	36	192	38	19	3	16

Fuente: Instituto Nacional de Estadística

9. Factores de Expansión

El factor de expansión es el inverso de la probabilidad de selección de la vivienda. El factor de expansión final, tiene el ajuste de la no-respuesta, considerando los resultados de incidencias de campo y el total poblacional proyectado para ese año, por esta razón los factores de expansión deben ser calculados una vez realizada la encuesta.

Las fórmulas con las que se calculará las probabilidades, vienen dadas por:

$$P(viv_{ijh}) = \left(\frac{A_h N_{jh}}{N_h} \right) \cdot \left(\frac{VE_{jh}}{VOE_{jh}} \right) = \left(\frac{A_h N_{jh}}{N_h} \right) \cdot \left(\frac{c}{VOE_{jh}} \right)$$

Donde:

- $P(viv_{ijh})$: Probabilidad de seleccionar la i – ésima vivienda de la j – ésima UPM, del estrato h de una UPM del área amanzanada.
- A_h : Número de UPM seleccionadas del estrato h .
- N_h : Número de viviendas del estrato h .
- N_{jh} : Número de viviendas en la j – ésima UPM del estrato h .
- VE_{jh} : Total de viviendas elegidas (viviendas elegidas ocupadas) en la j – ésima UPM, del estrato h , es decir, $VE_{jh} = c$
- VOE_{jh} : Total de viviendas objeto de estudio (viviendas ocupadas) listadas en la j – ésima UPM, del estrato h .
- c : Número fijo de viviendas seleccionadas en la última etapa, $c = 12$.

El procedimiento habitual indica que a los factores de expansión se aplique la corrección de las proyecciones de población, para que los mismos expresen el total de la población boliviana excepto las personas que habitan en viviendas colectivas que no son objeto de estudio en la encuesta.

10. Ajustes a los Factores Expansión

Por Incidencias Finales

Los factores de expansión se corrigen por incidencias de campo, es decir, que se utiliza los valores de la muestra efectiva por UPM.

El ajuste se realiza multiplicando el factor básico por la relación: $\frac{c}{V_{jh}}$

V_{jh} : es la muestra efectiva, número de viviendas con encuestas válidas en la j – ésima UPM del estrato h .

c : es el número de viviendas planificadas para ser encuestadas = 12

Por Proyecciones de Población al 2018

Para estimar el Total de hogares y el Total de población, se corrige el factor de expansión con las proyecciones de población que genere el INE para la gestión 2018, esta información por cada dominio de estudio.

La corrección se realiza multiplicando por: $\frac{P_h}{\hat{P}_h}$

Dónde:

P_h : Población en el h – ésimo dominio, según la proyección.

\hat{P}_h : Población en el h – ésimo dominio, a la que expande la encuesta.

Los factores de expansión finales estarán dados por:

$$F_{jh} = \frac{1}{P(Viv_{ijh})} \times \frac{c}{V_{jh}} \times \frac{P_h}{\hat{P}_h}$$

El resultado constituye el Factor de Expansión final que debe ser utilizado para la generación de indicadores.

11. Estimadores y Errores de Muestreo

Se ha señalado que el diseño empleado es complejo, por lo cual se requiere utilizar fórmulas especiales que consideran los efectos de la estratificación y la conglomeración al momento de obtener los resultados de la encuesta.

Para calcular los estimadores, es fundamental que la base de datos de la encuesta esté expandida, vale decir, que se empleen los factores de expansión. Para el cálculo de los intervalos de confianza, coeficiente de variación, error estándar, efecto de diseño, en otros, es necesario aplicar el diseño de la muestra antes de utilizar la base de datos en la generación de indicadores.

Se recomienda valerse de algún software estadístico que permita realizar el cálculo de los errores muestrales aplicando el diseño de muestra empleado.

12. Incidencia de campo

Una vez finalizado el operativo de campo se procede a realizar el ajuste de los coeficientes de variación de cada una de los departamentos, así como la incidencia de campo.

Tabla 8: Coeficiente de variación (CV) del tamaño de muestra ejecutada

Departamento/Ciudad		Número de viviendas (muestra ejecutada 2018)	Coeficiente de Variación esperada (CV) en base a la muestra ejecutada por departamento		
			CV de desempleo	CV de pobreza	CV de pobreza extrema
Nacional		11'195	2,7	1,4	2,0
Chuquisaca	Chuquisaca	619		4,4	5,6
La Paz	La Paz (Ciudad)	1046	7,1	6,5	11,4
	La Paz (Resto)	836		4,0	5,4
	La Paz (El Alto)	1063		5,1	7,9
Cochabamba	Cochabamba (Ciudad)	1424	6,1	4,0	6,7
	Cochabamba (Resto)	832		4,4	5,8
Oruro	Oruro	694		5,8	8,6
Potosí	Potosí	590		4,7	5,7
Tarija	Tarija	1016		4,7	7,4
Santa Cruz	Santa Cruz (Ciudad)	1190	8,1	6,0	12,0
	Santa Cruz (Resto)	785		5,9	9,0
Beni	Beni	653		5,3	7,4
Pando	Pando	447		7,6	10,8

Fuente: Instituto Nacional de Estadística

Según la definición de la Real Academia de la Lengua Española define incidencia como “acontecimiento que sobreviene en el curso de un asunto o negocio y tiene con él alguna conexión”. Esta definición genérica, en el ámbito de la investigación a través de encuestas, se refiere fundamentalmente a los acontecimientos que tienen lugar entre el número de muestra de viviendas planificadas para la recogida de información y lo que realmente se obtiene. Es decir, sería la diferencia entre lo planificado y lo ejecutado; entre la muestra teórica y la realmente conseguida.

A continuación se presenta la tabla 9, que refleja las incidencias de campo de la encuesta por Departamento, por ciudad capital, más la ciudad de El Alto, resto urbano y rural:

Tabla 9: Incidencia de campo por departamento

Muestra Planificada					Muestra Ejecutada				Incidencia de Campo (Tasa de Respuesta) en %			
	Total	Urbano	Resto Urbano	Rural	Total	Urbano	Resto Urbano	Rural	Total	Urbano	Resto Urbano	Rural
NACIONAL	11.448	6.396	2472	2580	11.195	6.200	2426	2569	97,8%	96,9%	98,1%	99,6%
Chuquisaca	624	276	72	276	619	272	71	276	99,2%	98,6%	98,6%	100,0%
La Paz	3.060	2.208	444	408	2.945	2.109	430	406	96,2%	95,5%	96,8%	99,5%
Cochabamba	2.292	1.452	468	372	2.256	1.424	461	371	98,4%	98,1%	98,5%	99,7%
Oruro	696	324	84	288	694	324	84	286	99,7%	100,0%	100,0%	99,3%
Potosí	600	120	156	324	590	117	153	320	98,3%	97,5%	98,1%	98,8%
Tarija	1.020	408	324	288	1.016	406	322	288	99,6%	99,5%	99,4%	100,0%
Santa Cruz	2.040	1.236	540	264	1.975	1.190	525	260	96,8%	96,3%	97,2%	98,5%
Beni	660	144	348	168	653	140	345	168	98,9%	97,2%	99,1%	100,0%
Pando	456	228	36	192	447	218	35	194	98,0%	95,6%	97,2%	101,0%

Fuente: Instituto Nacional de Estadística