

Diseño de muestra

Encuesta de Hogares 2012

03/12/2012

UNIDAD DE MUESTREO Y LOGISTICA

Contenido

INTRODUCCIÓN.....	4
ASPECTOS METODOLOGICOS DE LA ENCUESTA DE HOGARES 2012	4
1. Objetivos	4
2. Necesidad del rediseño de la muestra.....	4
3. Ámbito Geográfico	5
4. Universo de Estudio	5
5. Periodo de levantamiento de los datos	5
6. Características generales del diseño de muestra.	5
7. Marco Muestral.	6
8. Dominios de Estudio.	6
9. Unidades de Muestreo.	7
10. Selección de la muestra.	7
11. Cálculo del Tamaño de Muestra.	8
12. Distribución de la Muestra por Departamento y área geográfica	9
13. Criterios de Conglomeración y Estratificación del Marco de Muestreo	10
Variable de Conglomeración.....	10
Variable de Estratificación	11
14. Factores de Expansión	13
15. Ajustes a los Factores Expansión	14
16. Estimadores y Errores de Muestreo	15
17. Incidencia de campo	16

Contenido de tablas

Tabla 1: Estructura del Marco Muestral	6
Tabla 2: Coeficiente de variación de la muestra planificada	8
Tabla 3: Coeficiente de variación de la muestra ejecutada	9
Tabla 4: Distribución de la muestra de viviendas y UPMs por departamento y área geográfica	9
Tabla 5: Número de sectores por área geográfica	10
Tabla 6: Distribución de los conglomerados por departamento	11
Tabla 7: Estrato geográfico	11
Tabla 8: Estrato de pobreza	12
Tabla 9: Reagrupación del estrato geográfico	12
Tabla 10: Reagrupación del estrato de pobreza	12
Tabla 11: Incidencia de campo por departamento	16

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 retomaron 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, total rural.

Siendo a partir del 2011 en el que se realiza la desagregación a nivel total nacional, total urbano, total rural y departamental, y para el 2012 se ha previsto incluir las cuatro ciudades importantes de Bolivia (La Paz, Cochabamba, Santa Cruz y El Alto).

El documento consta de seis secciones la primera contiene los objetivos del diseño, segundo el universo de estudio, tercero características de la muestra, cuarto marco muestral, quinto dominios de estudio, sexto las unidades de la muestra, séptimo cálculo del tamaño de la muestra y última sección factores de expansión.

ASPECTOS METODOLOGICOS DE LA ENCUESTA DE HOGARES 2012

1. Objetivos

Los objetivos del presente diseño de muestra son:

- El objetivo principal es el generar indicadores relacionados con el empleo, desempleo y otras características económicas, que permitan 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 total urbano y total rural y a nivel departamental, en el caso de Beni y Pando se tendrán estimaciones conjuntas para ambos departamentos, y comparar con los datos obtenidos en la gestión 2011.
- Generar indicadores de ocupación para las ciudades capitales de los departamentos del eje central (La Paz, Cochabamba, Santa Cruz y El Alto).

2. Necesidad del rediseño de la muestra

Se estimó necesario replantear la Encuesta de Hogares en términos de un nuevo marco muestral (el Marco 2001) y un estudio profundo y extenso de estratificación, es así que a partir del 2011, la muestra de la EH incluye:

- a) una nueva estratificación del marco muestral
- b) la posibilidad de producir resultados a nivel departamento, excepto Beni y Pando, con resultados conjuntos.

3. Ámbito Geográfico

La encuesta de hogares 2012 cubre todos los departamentos del país: Chuquisaca, La Paz, Cochabamba, Oruro, Potosí, Tarija, Santa Cruz, Beni y Pando.

4. Universo de Estudio

La investigación está dirigida al conjunto de hogares establecidos en viviendas particulares ocupadas de las ciudades capitales, resto urbano y área rural de Bolivia, en el periodo definido para la ejecución de la encuesta en la gestión 2012.

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.

5. Periodo de levantamiento de los datos

El levantamiento de información primaria corresponde al periodo de octubre (16) a noviembre (15) de 2012.

6. Características generales del diseño de muestra.

EL diseño muestral aplicado combina la estratificación estadística y la conglomeración de áreas geográficas, con el muestreo multietápico, por tanto, puede ser caracterizado como:

- 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 2001, denominadas conglomerados;
- estratificado, porque previo a la selección de conglomerados, estos son clasificados en estratos con el propósito de mejorar la precisión de los estimadores y permitir un mejor control de la distribución de la muestra;
- multietápico, porque los elementos pertenecientes a la muestra se seleccionan en dos etapas en el área amanzanada; Etapa 1: Selección de conglomerados (UPM's); Etapa 2: selección de viviendas dentro la UPM elegida.

En área dispersa la selección se realiza en tres etapas; Etapa 1: Selección de conglomerados UPM's; Etapa 2: selección de segmento(s) censales dentro la UPM elegida; Etapa 3: selección de viviendas dentro del segmento(s) elegido.

7. Marco Muestral.

Se emplea el Marco Muestral construido en base a la información del último Censo Nacional de Población y Vivienda 2001(septiembre), siendo utilizado a partir del año 2002 en las encuestas a hogares.

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	Desdepto_first	Departamento: descripción y código
	Desprov_first	Provincia: descripción y código
	Dessecc_first	Sección municipal: descripción y código
Variables de organización municipal administrativa rural	Descant_first	Cantón: descripción y código
	Desciu_first	Comunidad: descripción y código
Variables de Organización Municipal administrativa urbana	Zona_first	Zona
	Imaz_manza_first	Manzano
Variables de organización Censal	Alocv_zona_first	Zona Censal
	Isect_sect	Sector Censal
	Iseg_segme_first	Segmento Censal
Variables para el diseño de muestra	Conglo	Conglomerados - Unidad Primaria de Muestreo - UPM
	Estratgeo Estrat_est_first	Estratos - Estrato geográfico - Estrato estadístico (NBI)
	Totviv_sum	Número de viviendas
	Tothom_hom_sum	Número de Hombres
	Totmuj_muj_sum	Número de Mujeres

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.

8. Dominios de Estudio.

Se garantiza generar estimaciones de indicadores de pobreza, con errores muestrales admisibles, para los siguientes niveles de desagregación: Nacional, Nacional Urbano, Nacional Rural y Departamental, excepto Beni y Pando, de los cuales se darán estimaciones de manera conjunta.

Para los indicadores de ocupación, los resultados podrán ser desagregados a nivel Nacional y ciudades capitales de los departamentos del eje central (La Paz, Cochabamba y Santa Cruz).

Realizar desagregaciones mayores de los resultados estará sujeto a la valoración de los errores muestrales.

Nivel de desagregación de la información

Generar indicadores que permitan conocer la evolución de la pobreza, el bienestar y la tenencia de las condiciones de vida de los hogares del país a nivel:

- a) Nacional
- b) Total Urbano
- c) Total Rural
- d) Departamental (En el caso de Beni y Pando se obtendrán estimaciones conjuntas)

Generar indicadores de ocupación para las siguientes ciudades capitales:

- a) La Paz
- b) Cochabamba
- c) Santa Cruz
- d) El Alto (La Paz)

9. Unidades de Muestreo.

Área Amanzanada

- UPM - La unidad primaria de muestreo es el conglomerado (UPM + Cantón + Sector Censal)
- USM - La unidad secundaria de muestreo conforman las viviendas particulares dentro de las UPM seleccionadas.

Área dispersa

- UPM - La unidad primaria de muestreo es el conglomerado (UPM + Cantón + Sector Censal)
- USM - La unidad secundaria de muestreo conforman los segmentos censales dentro de las UPM seleccionadas.
- UTM - La unidad terciaria de muestreo conforman las viviendas particulares dentro del segmento censal seleccionado.

10. Selección de la muestra.

Área Amanzanada

- En la primera etapa se utiliza la Selección Sistemática con probabilidad proporcional al tamaño (PPT) de viviendas.
- En la última etapa la selección es sistemática con arranque aleatorio

Área Dispersa

- En la primera y segunda etapa se utiliza la Selección Sistemática con probabilidad proporcional al tamaño (PPT) de viviendas.
- En la última etapa la selección es sistemática con arranque aleatorio.

11. 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 2007.

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. Se admite un error de muestreo máximo del 10% en estos estimadores.

Tabla 2: Coeficiente de variación de la muestra planificada

Departamento	Número de Viviendas (muestra planificada)	CV considerando tasas de la gestión 2007		
		CV de pobreza	CV de desempleo	CV de pobreza extrema
Chuquisaca	432	3.8	18.6	4.3
La Paz (ciudad)	1176	2.9	5.5	2.4
La Paz (no ciudad)	576	4.4	19.1	8.2
La Paz - El Alto	1008	3.3	5.4	4.2
Cochabamba (ciudad)	1188	3.3	5.7	3.1
Cochabamba (no ciudad)	672	4.4	11.3	8.3
Oruro	408	5.3	15.8	6.4
Potosí	576	3.0	16.2	3.2
Tarija	444	5.0	11.7	6.7
Santa Cruz (ciudad)	924	3.6	5.4	4.0
Santa Cruz (no ciudad)	708	4.0	10.1	8.0
Beni	336	5.8	9.4	8.6
Pando	288	7.8	16.2	9.2
Total	8736			

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

Tabla 3: Coeficiente de variación de la muestra ejecutada

Departamento	Número de Viviendas (muestra ejecutada)	CV considerando tasas de la gestión 2007		
		CV de pobreza	CV de desempleo	CV de pobreza extrema
Chuquisaca	432	3.8	18.5	4.3
La Paz (ciudad)	1135	2.9	5.6	2.5
La Paz (no ciudad)	575	4.4	19.0	8.2
La Paz - El Alto	977	3.3	5.5	4.3
Cochabamba (ciudad)	1112	3.4	5.9	3.2
Cochabamba (no ciudad)	668	4.5	11.4	8.4
Oruro	407	5.3	15.9	6.4
Potosí	574	3.0	16.2	3.2
Tarija	441	5.1	11.7	6.7
Santa Cruz (ciudad)	825	3.8	5.7	4.2
Santa Cruz (no ciudad)	692	4.0	10.2	8.1
Beni	325	5.9	9.5	8.8
Pando	252	8.4	17.2	9.8
Total	8415			

12. Distribución de la Muestra por Departamento y área geográfica

El cuadro siguiente presenta la distribución de la muestra por departamento y área geográfica, de la Encuesta de Hogares 2012:

Tabla 4: Distribución de la muestra de viviendas y UPMs por departamento y área geográfica

Departamento	Número de Viviendas de la muestra				Número de UPMs de la muestra			
	Total	Ciudad	Resto Urbano	Rural	Total	Ciudad	Resto Urbano	Rural
CHUQUISACA	432	144	24	264	36	12	2	22
LA PAZ	1752	1176	72	504	146	98	6	42
EL ALTO*	1008	1008	0	0	84	84	0	0
COCHABAMBA	1860	1188	192	480	155	99	16	40
ORURO	408	168	36	204	34	14	3	17
POTOSI	576	96	84	396	48	8	7	33
TARIJA	444	156	120	168	37	13	10	14

SANTA CRUZ	1632	924	336	372	136	77	28	31
BENI	336	72	156	108	28	6	13	9
PANDO	288	120	0	168	24	10	0	14
NACIONAL	8,736	5,052	1,020	2,664	728	421	85	222

* Para las ciudades capitales de los departamentos de La Paz, Cochabamba, Santa Cruz y El Alto se empleó el coeficiente de variación de la tasa de desempleo.

13. Criterios de Conglomeración y Estratificación del Marco de Muestreo

Variable de Conglomeración

La variable de conglomeración que el INE aplica a partir de la gestión 2002 hasta el 2009 en las Encuestas a Hogares, ha sido conformada con información y cartografía del Censo Nacional de Población y Vivienda del 2001. Los conglomerados (de aquí en adelante denominados como UPM, Unidad Primaria de Muestreo) se han agrupado en áreas geográficas de uno hasta cinco sectores censales, y así obtener un número similar de viviendas. El Marco Muestral utilizado hasta el 2009, está constituido por 16.790 conglomerados (UPMs).

Tal como muestra el cuadro de abajo, el 33.8% (5.682) de los conglomerados incluyen a más de un Sector Censal, lo cual implica que se incrementen las etapas de selección o que los segmentos seleccionados dentro una UPM se encuentren a bastante distancia ocasionando problemas en el operativo de campo.

Tabla 5: Número de sectores por área geográfica

		ÁREA					
		Urbano		Rural		Total	
		Recuento	% del N de la columna	Recuento	% del N de la columna	Recuento	% del N de la columna
Nro. de sectores	1	7947	65.1%	3161	68.9%	11108	66.2%
	2	4193	34.4%	1327	28.9%	5520	32.9%
	3	59	.5%	83	1.8%	142	.8%
	4	4	.0%	12	.3%	16	.1%
	5	0	.0%	4	.1%	4	.0%
	Total	12203	100.0%	4587	100.0%	16790	100.0%

Adicionalmente, se debe considerar que estas UPMs han cambiado considerablemente desde el censo 2001 a la fecha, la mayoría de los sectores del área amanzanada han crecido notablemente, mientras que algunas UPMs del área rural dispersa han disminuido de población, que para diferenciar a los sectores se toma en cuenta la variable de división organización municipal administrativa (cantón) la cual está inmersa en cada UPM, y previo a la variable de organización censal (sector censal).

Para evitar estas dificultades, se determinó tomar como conglomerado de primera etapa de selección de la muestra a los Sectores Censales.

Con esta modificación, el Marco Muestral para la Encuesta de Hogares 2011 está conformado por 22.658 conglomerados (UPM + Cantón + Sector Censal). Estos conglomerados contienen entre 40 y 120 viviendas en el área amanzanada y de 60 a 280 viviendas en el área dispersa. A continuación se presenta la distribución de los conglomerados en el país.

Tabla 6: Distribución de los conglomerados por departamento

Departamento	Nro. de Conglomerados (Sectores)	Porcentaje
Chuquisaca	1.121	4.9
La Paz	7.565	33.4
Cochabamba	3.914	17.3
Oruro	1.205	5.3
Potosí	1.743	7.7
Tarija	1.006	4.4
Santa Cruz	5.196	22.9
Beni	787	3.5
Pando	121	0.5
Total	22.658	100

Variable de Estratificación

Con el propósito de obtener precisión en los estimadores en la encuesta de Hogares 2012 se estructuró y combinó los conglomerados de selección en la primera etapa. Dentro de cada departamento, se ha utilizado la combinación de las variables estrato geográfico (1: Ciudades capitales incluyendo El Alto; 2: Ciudades intermedias con 10,000 y más habitantes; 3: Centros poblados de 2,000 hasta menos de 10,000 habitantes; 4: Localidades con características amanzanadas de 250 hasta menos de 2,000 habitantes; 5: Áreas dispersas con menos de 250 habitantes) y el estrato de pobreza NBI (1: Alto; 2: Medio Alto; 3: Medio Bajo; 4: Bajo). La diferencia radica en la utilización de estas variables como estratificación de manera independiente en las encuestas a hogares desde el año 2002 al 2007 y el año 2009, y a partir del 2011 se realiza una combinación de estas variables para conformar los estratos, conformando $3 \times 4 = 12$ estratos por departamento, haciendo un total de 96 estratos para los nueve departamentos.

Previo a la combinación se reagrupa el estrato geográfico de la siguiente forma:

Tabla 7: Estrato geográfico

Estrato Geográfico	Descripción	Estrato de Reagrupación
1	Ciudades Capitales más la ciudad de El Alto	1. Ciudad capital
2	Ciudades Intermedias con 10,000 y más habitantes	2. Resto Urbano
3	Centros Poblados de 2,000 a menos de 10,000 habitantes	
4	Localidades con características amanzanadas de 250 a menos de 2,000 habitantes	3. Rural
5	Áreas Dispersas de menos de 250 habitantes	

Tabla 8: Estrato de pobreza

Estrato de Pobreza NBI ¹	Descripción
1	Alto, son aquellas unidades muestrales que se encuentran entre el umbral y las necesidades básicas satisfechas.
2	Medio Alto, son unidades muestrales que están en el umbral de pobreza.
3	Medio Bajo, son unidades muestrales que están en la clasificación moderada de pobreza.
4	Bajo, son unidades muestrales que están entre la indigencia y marginalidad de pobreza.

Para la Encuesta de Hogares 2012 se considera la combinación de las mismas variables, realizando un ajuste de los estratos para evitar el excesivo número de ellos con el que se contaba, y corregir las variaciones de las estimaciones de los estratos, con el propósito de afinar o ajustar factores de expansión.

Para este fin, se utilizó los estratos de estratificación del diseño previo al levantamiento de campo en las encuestas a hogares, el de conglomeración de población (geográfico) y el de niveles de pobreza (NBI), previamente se procedió a la reagrupación de los estratos.

Tabla 9: Reagrupación del estrato geográfico

Estrato Geográfico	Descripción	Reagrupación geográfica
1	Ciudades Capitales más la ciudad de El Alto	1. Urbano
2	Ciudades Intermedias con 10,000 y más habitantes	
3	Centros Poblados de 2,000 a menos de 10,000 habitantes	
4	Localidades con características amezanadas de 250 a menos de 2.000 habitantes	2. Rural
5	Áreas Dispersas de menos de 250 habitantes	

Tabla 10: Reagrupación del estrato de pobreza

Estrato de Pobreza NBI	Descripción	Reagrupación de pobreza
1	Alto, son aquellas unidades muestrales que se encuentran entre el umbral y las necesidades básicas satisfechas.	1. Alto (No pobre)
2	Medio Alto, son unidades muestrales que están en el umbral de pobreza.	
3	Medio Bajo, son unidades muestrales que están en la clasificación moderada de pobreza.	2. Bajo (pobre)
4	Bajo, son unidades muestrales que están entre la indigencia y marginalidad de pobreza.	

¹ El estrato de pobreza, denominado estrato estadístico, es resultado de un tratamiento del indicador de Necesidades Básicas Insatisfechas (NBI), con el método cluster, mediante el cual se caracterizó las UPMs por nivel de pobreza

Mediante la combinación de las últimas columnas de los cuadros de arriba geográfica y pobreza se han conformado los estratos para cada departamento. Conformando $2 \times 2 = 4$ estratos por departamento, haciendo un total de 36 estratos para los nueve departamentos.

14. Factores de Expansión

El factor de expansión preliminar 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:

Área Amanzanada

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

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's 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, $c = 12$.

Área Dispersa

$$P(viv_{ikjh}) = \left(\frac{A_h N_{jh}}{N_h} \right) \cdot \left(\frac{B_{jh} N_{kjh}}{N'_{jh}} \right) \cdot \left(\frac{VE_{kjh}}{VOE_{kjh}} \right) = \left(\frac{A_h N_{jh}}{N_h} \right) \cdot \left(\frac{B_{jh} N_{kjh}}{N'_{jh}} \right) \cdot \left(\frac{c}{VOE_{kjh}} \right)$$

donde:

- $P(viv_{ikjh})$: Probabilidad de seleccionar la i-ésima vivienda de la k-ésima USM, de la j-ésima UPM, del estrato h en una UPM del área dispersa.

- A_h : Número de UPM's 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.
 B_{jh} : SM's (grupo de segmentos censales) seleccionados en la j-ésima UPM del estrato h, se tomará uno solo, entonces $B_{jh} = 1$.
 N_{kjh} : Número de viviendas en la k-ésima USM (grupo de segmentos censales), de la j-ésima UPM del estrato h.
 N'_{jh} : Número de viviendas en la j-ésima UPM del estrato h.
 c : Número fijo de viviendas seleccionadas en la última etapa, $c = 12$.
 VE_{kjh} : Total de viviendas elegidas (viviendas elegidas ocupadas) en la k-ésima USM (segmentos censales) de la j-ésima UPM, del estrato h, donde, $VE_{kjh} = c$.
 VOE_{kjh} : Total de viviendas objeto de estudio (viviendas ocupadas) listadas actualizadas en la k-ésima USM (grupo de segmentos censales) de la j-ésima UPM, del estrato h.

Tomando en cuenta que en el área dispersa no se realizará la actualización de toda la UPM elegida en la primera etapa, sino del segmento seleccionado, entonces los valores $N_{jh} = N'_{jh}$, por tanto se compensan.

Además, considerando que $B_{jh} = 1$, la fórmula del área dispersa queda de la siguiente manera:

$$P(viv_{ikjh}) = \left(\frac{A_h N_{kjh}}{N_h} \right) \cdot \left(\frac{c}{VOE_{kjh}} \right)$$

En consecuencia, para el cálculo de las probabilidades se opera como si fuera en dos etapas.

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.

15. 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 2011

Para estimar el Total de hogares y el Total de población, es necesario corregir el factor de expansión con las proyecciones de población que genere el INE para el año 2011, esta información se requiere por cada dominio de estudio.

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

Donde:

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:

Área Amanzanada

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

Área Dispersa

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

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

16. 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; entre los programas estadísticos más utilizados para realizar este procedimiento están el SPSS y el STATA.

Las variables que se utilizarán en el procedimiento son:

- a) El Factor de expansión calculado de la manera que se indica en el acápite anterior.
- b) La variable Conglomerado que se encuentra en la base de datos de la muestra y el Marco.
- c) La variable Estrato combinado existente en la base de datos de la muestra y el Marco.

17. Incidencia de campo

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 la planificación de la recogida de información y lo que realmente se obtiene. Es decir, sería la diferencia entre lo planificado y lo obtenido; entre la muestra teórica y la realmente conseguida.

La mayor parte de las incidencias ocurren en el momento en el que los entrevistadores tratan de localizar a las unidades a entrevistar, aquellas que han sido seleccionadas en el diseño muestral. En ese proceso de búsqueda se producen diversas situaciones que, en muchos casos, impiden la realización de la entrevista a las personas fijadas. Nos estamos refiriendo a problemas de localización del encuestado, falta de cooperación, rechazo, etc. Entonces, denominaríamos al término incidencias como: “el conjunto de estados finales en los que quedan clasificadas las unidades de la muestra cuyo cuestionario no ha sido recogido”. a continuación mostramos las incidencias de campo por departamento.

Tabla 11: Incidencia de campo por departamento

Departamento	Muestra planificada			Muestra Ejecutada			Porcentaje de incidencia de campo		
	Total	Urbano	Rural	Total	Urbano	Rural	Total	Urbano	Rural
CHUQUIS ACA	432	168	264	432	168	264	100.0%	100.0%	100.0%
LA PAZ	1752	1248	504	1710	1206	504	97.6%	96.6%	100.0%
EL ALTO*	1008	1008	0	977	977	0	96.9%	96.9%	---
COCHABAMBA	1860	1380	480	1780	1301	479	95.7%	94.3%	99.8%
ORURO	408	204	204	407	203	204	99.8%	99.5%	100.0%
POTOSI	576	180	396	574	179	395	99.7%	99.4%	99.7%
TARIJA	444	276	168	441	275	166	99.3%	99.6%	98.8%
SANTA CRUZ	1632	1260	372	1517	1152	365	93.0%	91.4%	98.1%
BENI	336	228	108	325	223	102	96.7%	97.8%	94.4%
PANDO	288	120	168	252	93	159	87.5%	77.5%	94.6%
Total	8,736	6072	2,664	8415	5777	2638	96.3%	95.1%	99.0%