

**DISEÑO MUESTRAL:
ENCUESTA AGROPECUARIA 2015**

**INSTITUTO NACIONAL DE ESTADISTICA
UNIDAD DE MUESTREO Y LOGÍSTICA**

**LA PAZ – BOLIVIA
2015**

1. Introducción

La Encuesta Agropecuaria (EA) 2015, que es llevada a cabo por el INE en conjunto con el MDRyT, se constituye como insumo de información, por lo cual tiene los objetivos de producir estimaciones de los diversos indicadores económicos, profundizando el conocimiento sobre el uso de la tierra, la conservación de los recursos naturales agropecuarios, soberanía ambiental con desarrollo integral y adaptación al cambio climático. Así mismo permitirá profundizar los resultados del censo agropecuario en cuanto a la producción y destino de la producción agrícola y pecuaria, indicadores zootécnicos, accesibilidad a créditos agropecuarios entre otros temas.

El presente documento describe el diseño muestral para la EA-2015, como tal su propósito es precisar la metodología que se ajusta a la información proporcionada por el CNA-2013 y el marco muestral generado a partir de los resultados de este censo. Es así que también se presentan aproximaciones y/o estimaciones para los errores muestrales, tasa de respuesta usados para la determinación de los tamaños de muestra, selección de la muestra y ponderaciones de diseño y ponderadores muestrales.

2. Antecedentes

El Instituto Nacional de Estadística, en base al marco que generó el II Censo Nacional Agropecuario de 1984, produjo información estadística del sector agropecuario a través de encuestas, las mismas se ejecutaron los años 1990, 1992, 1994 y 1997, respectivamente con representatividad de los nueve departamentos.

El 2008 se realizó la Encuesta Nacional Agropecuaria (ENA-2008), que capturó datos actualizados sobre superficie, producción y rendimiento para el año agrícola 2007-2008 además de la población pecuaria por especies a nivel departamental. El marco muestral para esta encuesta se generó a partir del Censo Nacional de Población y Vivienda 2001. Se planificó recolección de información para 1'062 UPMs en la primera etapa y más de 8'000 unidades UPAs para la segunda etapa.

Tabla 1.

Distribución de la muestra por departamento, con base al área rural dispersa del CNPV 2001, ENA-2008					
Departamento	Nro de municipios	Nro de viviendas	Nro de segmentos	Nro de segmentos* programados (muestra)	Nro de boletas programadas (muestra)
Chuquisaca	28	76'465	1'994	108	850
La Paz	75	236'552	6'271	266	2'008
Cochabamba	44	169'504	5'377	203	1'593
Oruro	34	47'881	1'133	70	560
Potosí	38	134'066	3'427	173	1'245
Tarja	11	33'799	941	49	379
Santa Cruz	50	85'627	2'582	134	968
Beni	19	19'642	752	42	307
Pando	15	5'929	224	17	112
TOTAL	314	809'465	22'701	1'062	8'022

(*) Un segmento = una UPM

Fuente: ENA-2008

Los resultados de la ENA-2008 permitieron al INE, en coordinación con el Ministerio de Desarrollo Rural y Tierras (MDRyT), establecer una única fuente de información del sector agropecuario, estructurada tanto a nivel departamental como nacional.

Después de 29 años se realizó el Censo Nacional Agropecuario (CNA) el cual ejecuto la captura de información entre septiembre a noviembre de 2013. A partir del Censo Nacional Agropecuario 2013 se pretende contar con información actualizada y estructural de las actividades agrícola, ganadera, forestal, caza, pesca y otras, a través de la cuantificación de las UPA. Para la ejecución del operativo de campo del Censo se utilizaron los datos de la Actualización Cartográfica multipropósito ACM, debido a la carencia de un Marco Agropecuario, el cual fue actualizado previo operativo de campo, con la finalidad de establecer los marcos de trabajo a nivel de las unidades geográficas necesarias, en todo el país y permitir así garantizar la cobertura a nivel nacional de todas las unidades censales "susceptibles" al relevamiento de la información deseada. En el CNA-2013 a partir de la conformación del marco geográfico, cambia la denominación de "Vivienda" por la de "Unidad Censal", donde estas, teóricamente, dentro del área rural/dispersa son potenciales Unidades Productoras Agropecuarias (UPA).

3. Objetivos del diseño muestral

3.1. Objetivo general

Implementar un diseño muestral eficiente mediante una metodología adecuada y acorde con los lineamientos para el cumplimiento de los propósitos y objetivo de la EA-2015 que son "Actualizar y profundizar la información agropecuaria para estimar la superficie, producción y destino agro-pecuario a nivel nacional y departamental a partir de los resultados del CNA-2013"¹.

3.2. Objetivos específicos

- Establecer el tamaño de la muestra, así como su distribución;
- Definición del marco muestral;
- Implementación del método de selección de la muestra;
- Cálculo de probabilidades de selección y cálculo de los factores de expansión (ponderadores muestrales).

4. Diseño muestral

El diseño muestral de la Encuesta Agropecuaria, responde y se adecua a la información que se desea obtener de acuerdo a los objetivos de la encuesta la cual es actualizar y profundizar la información estadística agropecuaria para estimar la superficie, producción y destino agrícola como también pecuario a nivel nacional y departamental, entre otros, a partir de los resultados del censo agropecuario.

4.1. Universo de estudio

¹ Perfil del proyecto "Encuesta Agropecuaria EA-2015".

Las unidades de investigación de la encuesta agropecuaria es la Unidad de Producción Agropecuaria (UPA), la cual se determina como todo terreno que se utiliza total o parcialmente para la producción agrícola, pecuaria o ambas, por un productor (solo o con la ayuda de otras personas) sin considerar el régimen de tenencia ni la condición jurídica, en el área rural de Bolivia. También se define la UPA como el conjunto de tierras y/o ganado o aves, que es explotada o aprovechada por una persona llamada Productor (se incluye, en el estudio también, a las UPAs que se constituyen como personas jurídicas), sin límite de tamaño de superficie o de número de animales.

4.2. Ámbito geográfico

La encuesta cubrirá el área rural de los nueve departamentos de Bolivia.

4.3. Periodo de recolección de la información

La planificación del operativo de campo para la recolección de la información comprende los meses calendario de octubre y noviembre de 2015.

4.4. Unidades de la encuesta

- **Unidades de observación.-** La unidad de producción agropecuaria (UPA).
- **Unidades de información.-** El productor² o informante calificado (e. g. la esposa, o esposo, hijos mayores de 12 años o administrador).
- **Unidades de muestreo.-** Para el caso de las UPAs que pertenecen a comunidad, se toma como unidades de muestreo la comunidad y sus UPAs. Por otro lado, para las UPAs que no pertenecen a comunidad, se toman como unidades de muestreo exclusivamente a las UPAs.
- **Unidades de la muestra.-** Son las comunidades y UPAs seleccionadas para la muestra.
- **Unidades de análisis.-** Estos se constituirán en las UPAs con los valores de la información relevada.

4.5. Tipo de muestreo

Dado que el universo de estudio está constituido por la unión excluyente de dos tipos de UPAs (UPAs que pertenecen a comunidad y UPAs que no pertenecen a comunidad), las cuales conformaron dos marcos muestrales, se implementa un diseño muestral que combina dos tipos de muestreo: un muestreo probabilístico bi-etápico y estratificado por conglomerado en primera etapa, para el marco muestral de comunidades; y un muestreo probabilístico estratificado, para el marco muestral de UPAs dispersas.

² Es la persona individual o colectiva, que tiene la iniciativa económica y técnica, y la responsabilidad del manejo de la UPA. El productor es quien toma las decisiones sobre la siembra, venta de las cosechas, crianza de animales, etc. Afronta los riesgos económicos y disfruta de los beneficios resultantes de su trabajo.

Adicionalmente, para el diseño muestral fue necesario considerar de manera especial a un tipo de UPAs dispersas que aportan de manera significativa al total de la producción departamental y nacional (i.e. tienen más producción que el resto de las población de UPAs dispersas). Con el propósito de captar información de estas UPAs, denominadas grandes, se reservó una parte de la muestra de manera exclusiva, para esto se buscaron, de entre el marco de las UPAs dispersas aquellas que tienen mayor producción agrícola o pecuaria o avícola registrada en el CNA-2013, llegando a seleccionar las más importantes las cuales son alrededor del 3% de todas las UPAs; el proceso de selección se realizó dentro cada uno de los nueve departamentos y como variables de selección se tomaron: los rendimientos o volúmenes de producción de los productos o especies agrícolas más representativos de cada departamento para el año agrícola 2012-2013; las unidades o cabezas de ganado de las especies de animales más representativas de cada departamento para el año 2013 (enero a noviembre); y las unidades de aves producidas para el año 2013. Este diseño para las UPAs dispersas se especifica como 'selección de inclusión forzosa' (o también de selección con probabilidad uno), el cual pretende captar información de todas estas UPAs.

El diseño de la encuesta tiene las siguientes características:

Probabilístico: Porque la unidad de selección (comunidades y/o UPAs), tienen una probabilidad conocida y distinta de cero de ser seleccionadas para formar parte de la muestra.

Por conglomerados: Debido a que las unidades objeto de estudio (UPAs de comunidad, excluyendo las UPAs dispersas) se encuentran agrupadas en áreas conformadas, con fines de muestreo, en base a la información y cartografía del CNA-2013, las cuales son las comunidades.

Estratificado: Se agrupan las unidades de muestreo (UPMs, para el caso del marco de comunidades; UPAs, para el caso de UPAs dispersas) en estratos, permitiendo mejorar la distribución de la muestra.

Bi-etápico: Porque se tiene dos etapas de selección (solo corresponde a las UPAs de comunidades), en la primera etapa se seleccionan comunidades (UPMs) y en la segunda etapa se seleccionan UPAs (USMs) al interior de la comunidad seleccionada.

Por inclusión forzosa: Son elegidas con antelación las UPAs dispersas que son consideradas como productoras importantes.

4.6. Dominios de estudio

Un factor que determina los dominios de diseños es el tamaño de muestra total, ya que para cada dominio se deberá calcular un tamaño de muestra adecuado en función a indicadores básicos con un número mínimo de unidades de muestra para lograr una precisión deseada, por tanto no es conveniente incrementar la cantidad de los dominios de estudio.

Se usa como primer factor determinante para la construcción de los dominios a las áreas rurales de la distribución política del territorio nacional en los nueve departamentos (Chuquisaca, La Paz, Cochabamba, Oruro, Potosí, Tarija, Santa Cruz, Beni y Pando). Para mejora esta primera formación de dominios se usa también la distribución del territorio en función a 'pisos ecológicos' o 'zonas agro-productivas' puesto que están muy relacionadas con la actividad agrícola o pecuaria de las comunidades y UPAs. Estas zonas tienen la característica que pueden estar presentes de manera múltiple en los territorios de diferentes departamentos, por tanto se establece usar siete dominios³, los cuales combinados con los departamentos, llegan a conformar 21 sub-dominios. Finalmente se tienen los siguientes dominios definidos para estimaciones de indicadores básicos de referencia a los diferentes niveles de precisión deseados.

- Nacional (rural).
- Departamental (rural), los nueve departamentos.
- Dominios por departamento y zona agro-productiva: Chuquisaca (chaco y valles); La Paz (altiplano, amazonia, valles y yungas-y-chapare); Cochabamba (altiplano, valles, yungas-y-chapare); Oruro (altiplano); Potosí (altiplano y valles); Tarija (altiplano, chaco y valles); Santa Cruz (chaco, chiquitania-y-pantanal y valles); Beni (amazonia y llanuras-y-sabanas); Pando (amazonia).

4.7. Nivel de desagregación de la información

La precisión deseada para cada una de las combinaciones entre departamento y zona agroproductiva (21 dominios), solo puede garantizarse para los indicadores de referencia con los que se calcula la muestra. Si se desea estudiar la información de manera desagregada para el resto de los indicadores o parámetros a estimar, se recomienda valorar la precisión en función al sesgo de las estimaciones respecto de los valores poblacionales, por tanto para el estudio de otras variables o indicadores es recomendable agregar la información a nivel departamental.

4.8. Selección de la muestra

Muestra de UPMs: Su selección es hecha de manera aleatoria e independiente, esto corresponde a la primera etapa y se aplica a las comunidades.

Muestra de UPAs de comunidad: En la segunda etapa de la muestra bi-etápica, se seleccionan las UPAs de manera sistemática y con arranque aleatorio.

Muestra de UPAs dispersas aleatorias: Su selección es irrestricta aleatoria.

Muestra de UPAs dispersas forzosas: Fueron seleccionadas UPAs específicas, esto es: selección con 'probabilidad uno'.

³ El ordenamiento según pisos ecológicos es tomado del CNA-2013, escogiéndose de entre otras clasificaciones: '5 zonas agroproductivas' y '19 zonas agroproductivas', siendo la más eficiente para el diseño la de '7 zonas agroproductivas'.

4.9. Desviaciones del diseño de la muestra

Esto se puede mostrar mediante el rendimiento de la encuesta y las incidencias externas relacionadas con la recolección de la información, para cada tipo de muestra de UPA.

Muestra de UPA de comunidad.- Los principales factores para no encontrar las unidades de la muestra se debieron a: conflictos sociales registrados, especialmente en el departamento de Beni; la poca existencia de unidades de observación u informantes, en la segunda etapa de muestreo.

Muestra de UPA que no pertenecen a comunidad.- La cobertura de estas UPA no fue total debido principalmente a la localización del informante, dadas las características de ubicación de las mismas UPA.

4.9.1. Tasa de respuesta (complemento de la tasa de no respuesta)

Las siguientes tasas reflejan la cobertura a nivel departamental, de boletas con información, para cada muestra según el tipo de UPA.

Bolivia: 99.8% (UPM), 99.7% (UPA de comunidad), 88.6% (UPA forzosas), 99.1% (UPA aleatorias).

Chuquisaca: 100% (UPM), 98.9% (UPA de comunidad), 90% (UPA forzosas), 100% (UPA aleatorias).

La Paz: 99.3% (UPM), 98.2% (UPA de comunidad), 92.7% (UPA forzosas), 97.3% (UPA aleatorias).

Cochabamba: 100% (UPM), 100% (UPA de comunidad), 95.3% (UPA forzosas), 100% (UPA aleatorias).

Oruro: 100% (UPM), 97.4% (UPA de comunidad), 90.9% (UPA forzosas), 100% (UPA aleatorias).

Potosí: 100% (UPM), 99.9% (UPA de comunidad), 100% (UPA forzosas), 100% (UPA aleatorias).

Tarija: 100% (UPM), 98.8% (UPA de comunidad), 84.2% (UPA forzosas), 95.5% (UPA aleatorias).

Santa Cruz: 100% (UPM), 98.9% (UPA de comunidad), 80.8% (UPA forzosas), 100% (UPA aleatorias).

Beni: 98.1% (UPM), 86.2% (UPA de comunidad), 88% (UPA forzosas), 100% (UPA aleatorias).

Pando: 100% (UPM), 94.7% (UPA de comunidad), 80% (UPA forzosas), 84.6% (UPA aleatorias).

4.9.2. Ponderación

Primero se calculan las ponderaciones de diseño las cuales son los recíprocos de las probabilidades de selección, para cada tipo de muestra de UPA.

UPAs de comunidad.-

Probabilidad de selección de primera etapa para la i-esima UPM.

$$P_{1hi} = n_h/N_h$$

Probabilidad de selección de segunda etapa para cada UPA en la i-esima UPM.

$$P_{2hi} = 9/L_{hi}$$

Donde:

n_h = número de UPMs seleccionadas en el estrato h ;

N_h = número total de UPMs en el estrato h ;

L_{hi} = número de UPAs que figuran en el listado de UPAs en la UPM i del estrato h .

La probabilidad de selección total es: $P_{hi} = P_{1hi} * P_{2hi}$; la ponderación de diseño será el recíproco de esta. A estas ponderaciones de diseño se les ajusta las tasas de respuestas de UPMs y UPAs a nivel de estrato, para así obtener las ponderaciones de muestreo.

UPAs dispersas aleatorias.-

Solo hay una etapa de selección, por tanto solo hay una probabilidad de selección.

$$P_{hi} = n_h/N_h$$

Donde:

n_h = número de UPAs seleccionadas en el estrato;

N_h = número total de UPAs en el estrato h .

La ponderación de diseño será el recíproco de esta, al resultado se les ajusta las tasas de respuestas

UPAs dispersas forzosas.-

Por su forma de selección su probabilidad de selección es uno, entonces sus ponderadores de diseño tienen el valor de *uno*.

4.9.3. Estimaciones de varianzas

Las estimaciones de varianzas para las estimaciones de totales se calculan por simple adición de las estimaciones de las varianzas obtenidos con cada uno de los diseños. Sin embargo como las UPA forzosas tienen probabilidad de selección 'uno', no tienen varianza muestral; estas UPA contribuyen a la precisión de las estimaciones reduciendo el coeficiente de variación global.