



Instructivo para el Cálculo de la Muestra – Formulario Comunidad

Equipo sectorial SIASAR Colombia

El objetivo de este documento describir las actividades, procesos y procedimientos

INSTRUCTIVO PARA EL MUESTREO DEL CUESTIONARIO COMUNIDAD SIASAR COLOMBIA

1. Introducción

Para desarrollar estudios donde sea necesario “la recolección, análisis e interpretación de datos, ya sea para ayudar en la toma de decisiones o para explicar condiciones regulares o irregulares de algún fenómeno o estudio aplicado, de ocurrencia en forma aleatoria o condicional”¹ en una población se debe utilizar un método estadístico que proporcione condiciones rigurosas en la captura y tratamiento de la información. Este manual contiene los métodos para determinar el tamaño de la muestra y la forma de selección de las viviendas a visitar en la zona rural dispersa y en áreas nucleadas para la obtención de la información del formulario de comunidad SIASAR, Anexo II.

2. Glosario

Población: Es el conjunto completo de individuos o elementos que presentan características comunes y observables. Para el Sistema de Información de Agua Potable y Saneamiento Rural, la población está compuesta por el número total de viviendas que compone una comunidad.

Muestra: “Es la parte de la población que efectivamente se mide, con el objeto de obtener información acerca de toda la población. La selección de la muestra se hace por un procedimiento que asegure en alta grado que sea representativa de la población²”

Censo. Es el proceso de observar la población completa³. Es decir, que la muestra es igual a la población.

Muestreo Aleatorio Simple. “Es un procedimiento de toma de muestra, en el que todas las muestras posibles, de un tamaño fijo, tienen igual probabilidad de ser seleccionada. Para obtener una muestra por este procedimiento, se enumeran todas las unidades muestrales, y se sortean unidades hasta completar el número requerido⁴”

¹ Bendita Mamani. 2012. P5

² Galbiati.2012. P3

³ Galbiati.2012. P4

⁴ Galbiati.2012. P5

Muestreo de conveniencia o por selección intencionada: Es una técnica de muestreo no probabilístico donde los sujetos son seleccionados dada la conveniente accesibilidad y proximidad de los sujetos para el investigador.

Viviendas nucleadas: centros poblados rurales (Declarados en el POT – PBOT – EOT), o Conjunto de viviendas próximas entre sí, con infraestructura colectiva para los servicios públicos, es decir, son aquellas comunidades rurales que presentan condiciones similares a comunidades urbanas. Una de las características más frecuentes es la proximidad de las viviendas, las cuales están habitualmente juntas o ubicadas a una distancia no mayor a mil metros.

Vivienda Rural Dispersa: Aquella donde las casas se encuentran situadas a distancias considerables una de la otra.

3. INSTRUCTIVO PARA COMUNIDADES UBICADAS EN VIVIENDAS NUCLEADAS

Para aplicar el modelo muestral se debe tener en cuenta el tamaño de la población.

- Para poblaciones de 1 a 50 viviendas se realizará censo para garantizar la representatividad de la información obtenida, por tal razón se debe visitar el total de las viviendas de la comunidad.

1 a 50 ➔ **Muestra = Población**

- Para poblaciones con más de 50 viviendas se utilizará la técnica de muestreo aleatorio simple.

>50 a 500 ➔ **Muestra = Muestreo aleatorio simple**

Para este último caso se aplicará la siguiente fórmula:

$$n = \frac{Z^2 p \times q \times N}{Ne^2 + Z^2 p \times q}$$


n = Tamaño de la muestra
 N = Tamaño de la población
 p = Probabilidad a favor
 q = Probabilidad en contra
 e = Error de estimación
 Z^5 = Nivel de confianza

Se asume para el levantamiento de información SIASAR un margen de error del 10% y un nivel de confianza del 90%, donde Z es igual a 1,64.

Nivel de Confianza	99%	98%	96%	95%	90%	80%	68%	50%
Z	2,58	2,33	2,05	1,96	1,64	1,28	1	0,67

Asimismo, cuando no contemos con suficiente información sobre la probabilidad de ocurrencia del evento se sugiere asignar a p y q un valor de 0.5, respectivamente.

Para poblaciones de tamaño mayor a 1000 viviendas se realizará muestreo por selección intencionada, donde por cada 100 viviendas existentes se tomará 4 de muestra.

>1000  **Muestreo por selección**
Muestra = intencionada (4 visitas por
cada 100 casas)

Ahora bien, para seleccionar el intervalo de viviendas que se deben visitar se empleará la siguiente fórmula.

⁵ El análisis se fundamenta en la distribución normal del estimador, teniendo en cuenta que aunque la población no sea normal, los estimadores que se utilizan normalmente sí tienden a serlo al tratarse de suma de realizaciones de la distribución.

>50 $k = \frac{N}{n}$

Dónde:

K = Intervalo
N= Población
n = Muestra

Ejemplo: Se realizará el levantamiento de información para SIASAR en el centro poblado de La Palma, que está compuesto por 236 viviendas.

$$n = \frac{(1,966)^2(0.5x0.5x236)}{((236)(0,05)^2+(1.96)^2(0.5x0.5))} = 146$$

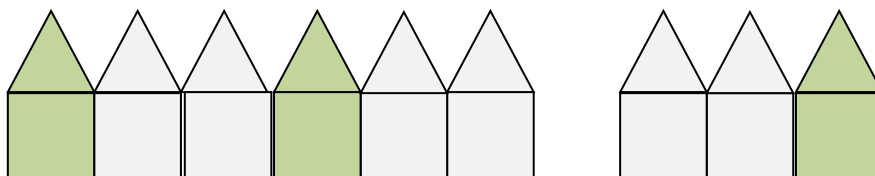
Por tanto, se deben realizar visita a 146 viviendas del centro poblado La Palma y aplicar el anexo II del formulario de comunidad.

Intervalo

$$k = \frac{236}{146} = 1,6$$

Es decir, que se deben realizar las visitas en intervalos de dos casas aproximadamente, como se observa en el siguiente diagrama 1.

Diagrama 1. Intervalo de visitas en viviendas




4. INSTRUCTIVO PARA COMUNIDADES UBICADAS EN VIVIENDAS DISPERSAS

Para determinar la proporción de la muestra se debe tener en cuenta el tamaño de la población de estudio.

Si la población es menor a 20 viviendas se debe realizar un censo, es decir que se debe realizar visita a la totalidad de las viviendas

≤ 20  **Muestra = Población**

Para poblaciones de tamaño entre 20 y 50 viviendas se realizará muestreo por selección intencionada, donde por cada 10 viviendas existentes se tomará 3 de muestra.

>20 a 50  **Muestreo por selección intencionada (3 visitas por cada 10 casas)**

Ejemplo. La vereda Agua Fría tiene una población dispersa de 38 viviendas. Se usa muestreo intencionado para determinar el tamaño de la muestra cómo se indica en la siguiente ecuación

$$n = \frac{38}{10} \times 3 = 11,4$$

Cuando el tamaño de la población está en un rango igual o mayor 50 hasta 1000 casas se desarrolla muestreo por selección intencionada, donde se selecciona 1 vivienda por cada 10 casas.

≥ 5 a 1000  **Muestra :** **Muestreo por selección intencionada (1 visitas por cada 10 casas)**


Ejemplo. La vereda de Monte Verde tiene una población de 385 casas, para determinar el número de intervenciones en las viviendas se emplea el método de muestreo por selección intencionada,

Donde

$$n = \frac{385}{10} \times 1 = 38.5$$

Por tanto, el tamaño de la muestra es de 39 casas y este será el número de visitas que se realizaran en la comunidad de la vereda Monte Verde.

Para poblaciones de tamaño mayor a 1000 viviendas se realizará muestreo por selección intencionada, donde por cada 100 viviendas existentes se tomara 4 de muestra.

>1000  **Muestra =** **Muestreo por selección intencionada (4 visitas por cada 100 casas)**

5. FACTOR DE AJUSTE SIASAR

Para hacer el reporte de datos de comunidad en la plataforma del Sistema de Información de Agua Potable y Saneamiento Rural SIASAR, se debe realizar el ajuste de proporcionalidad de los valores obtenidos en campo, por medio del tipo de muestreo que se aplicó para el levantamiento de la información.

Se debe aplicar el factor de proporción a los datos recolectado en campo, asumiendo que la muestra posee los mismos atributos que la población estudiada, por tanto, se usara la siguiente relación para determinar el dato que debe reportarse en la plataforma.

$$\text{Dato del reporte} = (B2\dots n / B1.4) * A4$$

B 2 = Cantidad de viviendas que representan la porción de la muestra en estudio.

B1.4 = muestra

A4 = Población

El cálculo se puede realizar en campo o en el momento de reportar los datos de manera manual o con la hoja de cálculo adjunta.

El factor de proporcionalidad debe aplicarse a los siguientes bloques del cuestionario de comunidad.

- B 2. Saneamiento. Existencia de infraestructura para las viviendas
- B 3. Saneamiento. Uso de la infraestructura MEJORADA de tipo 1 y tipo 2
- B 4. Saneamiento. Uso de la infraestructura NO MEJORADA
- Saneamiento. No Tienen infraestructura
- B5. Higiene en viviendas
- B6. Recolección y tratamiento de basuras

1. Ajuste manual en campo del factor de proporción.

Saneamiento. Existencia de infraestructura para las viviendas		Muestra	Total Viviendas (B...n/B1.4)*A4
Tipo de infraestructura			
Cantidad de viviendas que TIENEN infraestructura de saneamiento PROPIA mejorada tipo 1:		18	
<ul style="list-style-type: none"> • Descarga hidráulica (automática o manual) a red de alcantarillado o Tanque séptico 			
Cantidad de viviendas que TIENEN infraestructura de saneamiento PROPIA mejorada tipo 2:		29	
<ul style="list-style-type: none"> • Letrina de pozo mejorada con ventilación (VIP) • Letrina de pozo con losa • Letrina/inodoro de compostaje 			
Cantidad de viviendas que TIENEN OTRA infraestructura de saneamiento PROPIA del tipo no mejorada:		12	
<ul style="list-style-type: none"> • Descarga hidráulica (automática o manual) a cualquier otro lugar (la calle, patio o terreno, desagüe abierto, trinchera, drenaje abierto u otra ubicación) • Letrina de pozo sin losa o pozo abierto, de balde o colgante 			
Cantidad de viviendas que COMPARTEN infraestructura de saneamiento mejorada tipo 1:		5	
<ul style="list-style-type: none"> • Descarga hidráulica (automática o manual) a red de alcantarillado o Tanque séptico 			
Cantidad de viviendas que COMPARTEN infraestructura de saneamiento mejorada tipo 2:		4	
<ul style="list-style-type: none"> • Letrina de pozo mejorada con ventilación (VIP) • Letrina de pozo con losa • Letrina/inodoro de compostaje 			
Cantidad de viviendas que NO TIENEN infraestructura de saneamiento:		2	
<ul style="list-style-type: none"> • Defecación a campo abierto. 			
TOTAL		70	

En los formularios de campo de comunidad en los bloques B2, B3, B4, B5 y B6 contienen dos columnas; la primera contiene la información recolectada en campo, en la siguiente columna se Reporte el factor de ajuste proporción, Estas celdas representan los valores que deben suministrarse en la plataforma de SIASAR con el ajuste de proporción, usando la relación, (Dato del reporte = $(B_n/n)*N$).

Ejemplo: La vereda Lagunetas baja del municipio de Vélez del departamento Santander, cuenta con una comunidad compuesta por 85 viviendas; se determinó por medio del modelo muestral que se realizarán 70 visitas, donde se aplicará el anexo dos de comunidad en intervalos de una casa.

Los datos que obtuvieron son los siguientes:

Saneamiento. Existencia de infraestructura para las viviendas		
Tipo de Infraestructura	Muestra	Total Viviendas (B...n/B1.4)*A4)
B 2 Cantidad de viviendas que TIENEN infraestructura de saneamiento PROPIA mejorada tipo 1: • Descarga hidráulica (automática o manual) a Red de alcantarillado o Tanque séptico	18	
Cantidad de viviendas que TIENEN infraestructura de saneamiento PROPIA mejorada tipo 2: • Letrina de pozo mejorada con ventilación (VIP) • Letrina de pozo con losa • Letrina/inodoro de compostaje	29	
Cantidad de viviendas que TIENEN OTRA infraestructura de saneamiento PROPIA del tipo no mejorada: • Descarga hidráulica (automática o manual) a cualquier otro lugar (la calle, patio o terreno, desagüe abierto, trinchera, drenaje abierto u otra ubicación) • Letrina de pozo sin losa o pozo abierto, de balde o colgante	12	
Cantidad de viviendas que COMPARTEN infraestructura de saneamiento mejorada tipo 1: Descarga hidráulica (automática o manual) a Red de alcantarillado o Tanque séptico	5	
Cantidad de viviendas que COMPARTEN infraestructura de saneamiento mejorada tipo 2: • Letrina de pozo mejorada con ventilación (VIP) • Letrina de pozo con losa Letrina/inodoro de compostaje	4	
Cantidad de viviendas que NO TIENEN infraestructura de saneamiento: • Defecación a campo abierto.	2	
TOTAL	70	

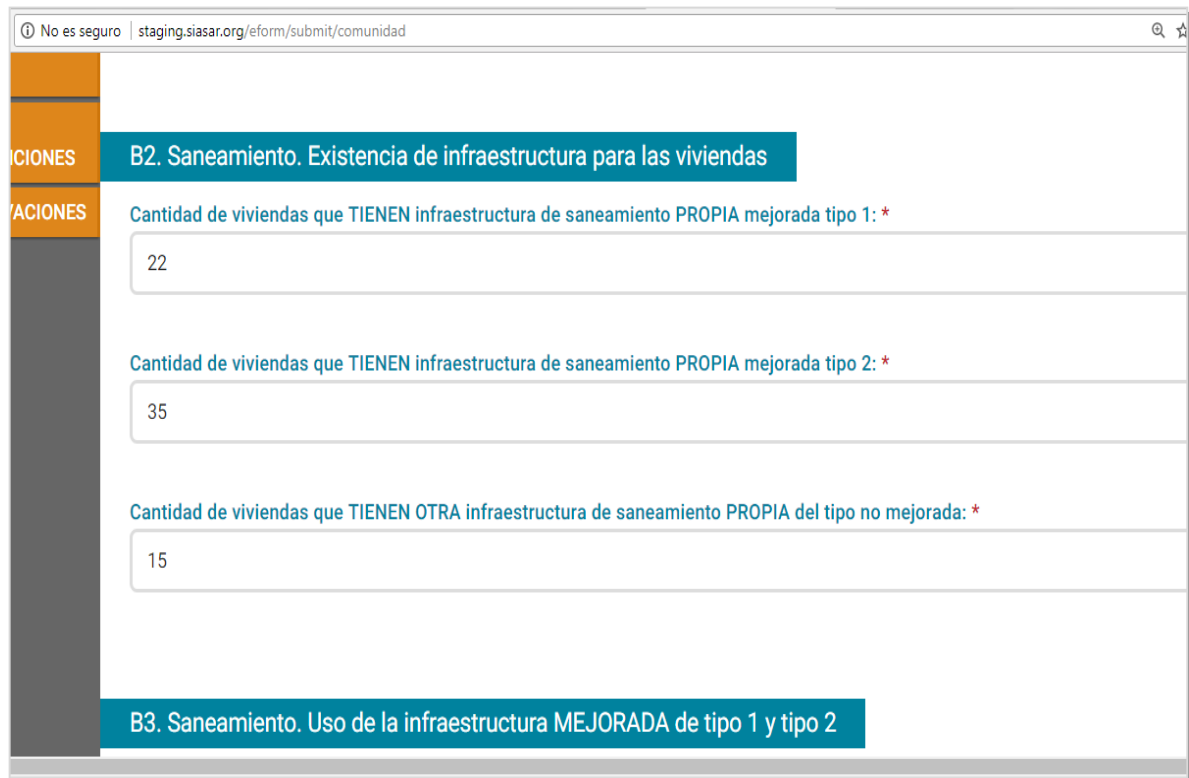
Para hallar el dato de reporte para plataforma, se dividen los de la Columna B2.1...n, con el total de viandas de la celda B1.4 que corresponde a la muestra y el producto de la anterior operación se multiplica por la celda A.4 (número total de viviendas).

Dato del reporte = $(18/70) * 85 = 22$

Saneamiento. Existencia de infraestructura para las viviendas		
Tipo de Infraestructura	Muestra	Total Viviendas (B...n/B1.4)*A4)
Cantidad de viviendas que TIENEN infraestructura de saneamiento PROPIA mejorada tipo 1: <ul style="list-style-type: none"> Descarga hidráulica (automática o manual) a Red de alcantarillado o Tanque séptico 	18	22
B 2 Cantidad de viviendas que TIENEN infraestructura de saneamiento PROPIA mejorada tipo 2: <ul style="list-style-type: none"> Letrina de pozo mejorada con ventilación (VIP) Letrina de pozo con losa Letrina/inodoro de conectarse 	29	

2. Reporte en la plataforma SIASAR.

El valor que se reporta en la plataforma es el dato de la columna DP, tal como se muestra en la siguiente imagen.



No es seguro | staging.siasar.org/eform/submit/comunidad

B2. Saneamiento. Existencia de infraestructura para las viviendas

Cantidad de viviendas que TIENEN infraestructura de saneamiento PROPIA mejorada tipo 1: *

22

Cantidad de viviendas que TIENEN infraestructura de saneamiento PROPIA mejorada tipo 2: *

35

Cantidad de viviendas que TIENEN OTRA infraestructura de saneamiento PROPIA del tipo no mejorada: *

15

B3. Saneamiento. Uso de la infraestructura MEJORADA de tipo 1 y tipo 2

Aplique este mismo procedimiento para los bloques B3, B4, B5 y B6

B2. Saneamiento. Existencia de infraestructuras en viviendas										
DATOS DE CAMPO							FACTOR DE AJUSTE SIASAR			
Código	Comunidad	Numero de viviendas totales	Número de la muestra	Cantidad de viviendas que TIENEN infraestructura de saneamiento PROPIA mejorada tipo 1	Cantidad de viviendas que TIENEN infraestructura de saneamiento PROPIA mejorada tipo 2	Cantidad de viviendas que TIENEN OTRA infraestructura de saneamiento PROPIA del tipo no mejorada:	Cantidad de viviendas que TIENEN infraestructura de saneamiento PROPIA mejorada tipo 1	Cantidad de viviendas que TIENEN infraestructura de saneamiento PROPIA mejorada tipo 2	Cantidad de viviendas que TIENEN OTRA infraestructura de saneamiento PROPIA del tipo no mejorada:	
1	Comunidad vereda lagunetas /Vélez/Santander	85	70	18	29	12	22	35	15	
5		85	70	24	8	7	29	10	9	
6		0	0	0	0	0	#DIV/0!	#DIV/0!	#DIV/0!	
7		0	0	0	0	0	#DIV/0!	#DIV/0!	#DIV/0!	
8		0	0	0	0	0	#DIV/0!	#DIV/0!	#DIV/0!	
9		0	0	0	0	0	#DIV/0!	#DIV/0!	#DIV/0!	
10		0	0	0	0	0	#DIV/0!	#DIV/0!	#DIV/0!	
11		0	0	0	0	0	#DIV/0!	#DIV/0!	#DIV/0!	

3. Ajuste en oficina del factor de proporción.

Este método consiste en la aplicación del anexo 1 de este instructivo, el cual es una modelo en Excel donde se suministran los datos de comunidad, población y muestra, la información recopilada en campo y el de manera automática calcula el factor de proporción.

El anexo contiene una hoja de cálculo para cada uno de los bloques en que requiere hacer el ajuste de relación de proporción entre muestra y población.