

Notas Técnicas y Anexos

Documento Metodológico

CONVENIO INTERADMINISTRATIVO UPME No. CV-008-202, IPSE No.147-2021,
SUSCRITO ENTRE LA UNIDAD DE PLANEACIÓN MINERO ENERGÉTICA- UPME; EL
INSTITUTO DE PLANIFICACIÓN Y PROMOCIÓN DE SOLUCIONES ENERGÉTICAS PARA LAS
ZONAS NO INTERCONECTADAS - IPSE Y LA UNIVERSIDAD DEL CAUCA.



Universidad
del Cauca
Vigilada Mineducación



PERS Cauca
Plan de energización Rural Sostenible

1 Introducción

El presente documento tiene como propósito básico explicar la metodología utilizada en el proceso de determinación de la muestra representativa para la población rural del Departamento del Cauca. Esto se realiza a partir del diseño de instrumentos estadísticos apropiados para el levantamiento de información primaria. Este documento será la guía para el trabajo de campo que se desarrollará durante la ejecución del Plan de Energización Rural y Sostenible (PERS) Cauca y busca tener información en torno al aprovechamiento de la energía por fuente, tenencia y usos para diversos sectores, en especial, el residencial y otros como el comercial e industrial.

Adicionalmente se cuenta con un repositorio GIT-Hub en el cuál se tienen los controles de versiones y archivos necesarios para el desarrollo del proyecto. (see https://github.com/cristian42253/PERS_Cauca)

2 Material adicional

Para realizar los cálculos de los factores de expansión, es de recordar que la aproximación es mediante la obtención de las probabilidades de que cada elemento de la encuesta represente la población objeto de estudio. En el caso particular del PERS Cauca se tienen las probabilidades desagregadas por municipios y también la agregación por sub-regiones

2.1 Cálculo factores de expansión

Recordemos que para el cálculo de los factores de expansión se utilizará las probabilidades de obtener cada uno de los elementos de la población objeto de estudio, para ello se usa la fórmula de Bayes desglosada en términos de las condiciones.

2.1.1 Formula de Bayes

De acuerdo al Teorema de Bayes, sabemos que una representación válida es,

$$P(\theta|x) = \frac{P(x|\theta)P(\theta)}{\int_{\Theta} P(x|\theta)P(\theta)d\theta}$$

Donde x corresponde de manera simple a la información con la que se cuenta, para el caso particular del estudio las viviendas del departamento del Cauca, con acceso a energía eléctrica y sin acceso a la misma según la fuente del DANE y UPME. θ es la información o parámetro que deseamos estimar, en nuestro caso particular la probabilidad de cada elemento de la población objeto de estudio para representar por medio de la encuesta. Finalmente $\int_{\Theta} P(x|\theta)P(\theta)d\theta$ representa la probabilidad total en el universo de estudio, en nuestro caso particular las viviendas de la zona rural del departamento del Cauca. Recordemos que según la visión del teorema de Bayes, el cálculo de las probabilidades se usa bajo el supuesto de información completa y fiable.

2.1.2 Aplicación Casos PERS- Cauca

Recordemos que los Factores de expansión para el PERS Cauca son:

Región	Número de Encuestas	Factor de Expansión	Población Expandida
Centro	507	168	85.641
Norte	306	266	81.139
Oriente	150	144	21.562
Occidente	707	27	19.227
Sur	206	217	44.748
Macizo	106	230	24.403
Bota caucana	109	82	8.923
Total	2091	**	285.643

Como nuestro objetivo es la obtención de una muestra representativa del departamento del Cauca visto a través de las subregiones que componen la división territorial del mismo, se asumirá que la escogencia de cada Subregión es equiprobable, lo que quiere decir que cada una de las viviendas al interior de las sub-regiones y de los municipios que cumplen con las condiciones propias del marco muestral tienen igual probabilidad de ser elegidas, de igual forma en conjunto con los datos del tamaño de muestral definido apriori se tiene la siguiente representación. $P(Viviendas \cap No \text{ Energía})$. Por tanto necesitaremos la probabilidad de obtener una Vivienda de la zona rural y con un IR mayor a 40%, dado que dichas viviendas no cuentan con fluido eléctrico. Para ello se define el cálculo de la siguiente forma.

$$P(Viviendas|No \text{ Energía}) = \frac{P(Viviendas \cap No \text{ Energía})}{P(Viviendas)}$$

Donde podemos obtener la probabilidad de obtener una vivienda que no cuenta con fluido electrico como,

$$P(Viviendas|No \text{ Energía}) * P(Viviendas) = P(Viviendas \cap No \text{ Energía})$$

Para el caso del centro se tiene

$$P(Viviendas_{Centro} \cap No \text{ Energía}_{Centro}) = \frac{1}{7} * \frac{1}{24}$$

en donde se tiene que para obtener el factor de expansión será en resumen el cociente entre 1 y la probabilidad de $P(Viviendas_{centro} \cap No \text{ Energía}_{centro})$. Para la zona centro, será $\frac{1}{0.00595238} = 768$.

Ahora bien, recordemos que para escoger un elemento de la población muestral en la sub región centro se tiene.

- $P(Sub - region_{Centro}) = \frac{1}{7}$, dado que es de interés capturar la máxima información posible del departamento, para ello se decidió que todas las sub-regiones estarían representadas (7 sub- regiones).
- $P(Intra Sub - region_{Centro}) = \frac{85641}{285643} * \frac{1}{7.1956} \approx \frac{1}{24}$, se multiplica la probabilidad de obtener una vivienda dentro la sub región por la proporción equivalente de la submuestra con las condiciones de no contar con el fluido eléctrico. (Intra sub región significa que se considera la población dentro de cada sub región)

De esta forma se tendrá que la probabilidad de cada escoger una vivienda con las características mencionadas anteriormente en la zona centro es la multiplicación entre ambas probabilidades. Es decir,

$$P(Viviendas_{centro} \cap No \text{ Energía}_{centro}) = \frac{1}{7} * \frac{1}{24}$$

Por lo que el Factor de Expansión (FE) será,

$$FE = \frac{1}{P(Viviendas_{centro} \cap No \text{ Energía}_{centro})} = 768$$

2.1.3 Casos completos

Tal y como se mostró anteriormente se puede notar que al realizar los procedimientos se tienen los valores correspondientes de las probavilidades sigueiendo el teorema de Bayes, lo cuaál permite obtener los factores de expansión, en resumen se puede ver los cálculos necesarios para obtener dichos resultados tal y como lo vemos a continuación.

	P(Viviendas n No energía)	P(Sub Región)	P(Intra Sub Región)	FE= (1/P(Viviendas n No energía))
Centro	$(1/7)*(1/24)$	1/7	$(85641/285643)*(1/7.1956)$	1056
Norte	$(1/7)*(1/38)$	1/7	$(81139/285643)*(1/10.2642)$	764
Oriente	$(1/7)*(7/144)$	1/7	$(21562/285643)*(1/1.55285)$	245
Occidente	$(1/7)*(7/27)$	1/7	$(19227/285643)*(1/0.2527)$	27
Sur	$(1/7)*(1/31)$	1/7	$(44748/285643)*(1/4.86396)$	396
Macizo	$(1/7)*(7/230)$	1/7	$(24403/285643)*(1/2.82371)$	371
Bota Caucana	$(1/7)*(7/82)$	1/7	$(8923/285643)*(1/0.365924)$	924

3 Anexo Obtención Muestra

A continuación se presentan los códigos por los cuales se obtuvieron las diferentes muestras para aplicar las encuestas.

```
## packages necesaries
packages = c("tidyverse", "pwr",
             "Sample.Size", "sampling","samplesize")

## Now load or install&load all
package.check <- lapply(
  packages,
  FUN = function(x) {
    if (!require(x, character.only = TRUE)) {
      install.packages(x, dependencies = TRUE)
      library(x, character.only = TRUE)
    }
  }
)

# Load packages
load_pkg <- rlang::quos(tidyverse, pwr, Sample.Size, sampling, samplesize)

invisible(lapply(lapply(load_pkg, rlang::quo_name),
                  library,
                  character.only = TRUE
                ))

# Calculate sampling
delta <- 1.45
sigma <- 14.5
d <- delta/sigma
pwr.t.test(d=d, sig.level=.05, power = .90, type = 'two.sample')
```