|  |  |
| --- | --- |
|  |  |

****

**CUARTO ESTUDIO NACIONAL DE SALUD BUCAL – ENSAB IV**

**MANUAL DE PROCEDIMIENTO DE ESTIMACIÓN**

**ARD-434- DOC 26**

**Bogotá,**

**Diciembre de 2014**



TABLA DE CONTENIDO

[1. MACRO ESTIMACIÓN IV ENSAB 3](#_Toc406062187)

[1.1. %ESTIMA\_RAZON 3](#_Toc406062188)

[1.1.1. PARÁMETROS 3](#_Toc406062189)

[1.1.2. USO 4](#_Toc406062190)

[1.1.3. RESULTADOS 5](#_Toc406062191)

# MACRO ESTIMACIÓN IV ENSAB

Para la estimación de resultados nacionales y agregados provenientes de la ENSAB IV se diseño una macro en el programas SAS. Teniendo en cuenta los diferentes resultados generados a partir de la información recolectada en campo para el estudio se generaron dos macros una para la estimación de porcentajes y promedios y la segunda para la estimación de porcentajes y totales estimados de población. A continuación se describen los parámetros de las funciones macro.

## %ESTIMA\_RAZON

La macro denominada ***%ESTIMA\_RAZON*** está diseñada para estimar porcentajes de población y razones de totales, de manera general con esta macro se pueden estimar promedios. Para la estimación de la incertidumbre asociada al diseño muestral, se utiliza linearización de taylor en conjunto con la especificación del diseño muestral.

### PARÁMETROS

* **CampoCol**: campo en el data set que trae la variable que define las categorías columna para el cuadro, debe ser una variable que este codificada de manera numérica. Para cada una de las categorías de campoCol se calculan distribuciones por los valores de campoFil de manera que totalice 100% por cada categoría de campoCol.
* **CampoFil**: campo en el data set que trae la variable que define las categorías fila para el cuadro, debe ser una variable que este codificada de manera numérica .
* **yk**: campo que define el numerador para el cálculo de la razón, numérica.
* **zk**: campo que define el denominador para el cálculo de la razón, númerica.
* **zkd**: campo que define el denominador para el cálculo de los porcentajes en el caso del cálculo de estimaciones para dominios, sólo toma valores 1 o 0.
* **BaseIn**: data set con los datos de entrada, las variables campoFil, campoCol, yk, zk, zkd y variables que definen el diseño.
* **BaseOut**: data set que se crea la macro y guarda los resultados.
* **Estrato**: lista de campos en el data set, separados por espacios, que definen los estratos del diseño muestral de cada una de las etapas consideradas.
* **UM**: lista de campos en el data set, separados por espacios, que definen los identificadores en cada una de las etapas del diseño muestral considerado.
* **Diseno**: lista de campos en el data set, separados por espacios, que definen los métodos de selección usados en cada una de las etapas del diseño muestral considerado.
* **N**: lista de campos en el data set, separados por espacios, que definen las cantidades totales de unidades de muestreo en el marco (N) de cada una de las etapas del diseño muestral consideradas.
* **nm**: lista de campos en el data set, separados por espacios, que definen la cantidad de unidades muestrales seleccionadas (nm) de cada una de las etapas del diseño muestral consideradas.
* **p**: lista de campos en el data set, separados por espacios, que definen las probabilidades de inclusión o selección de cada una de las etapas del diseño muestral consideradas.
* **fkFinal**: factor de expansión final a usar para el cálculo de los estimadores.
* **alphaSig**: nivel de significancia usado para el cálculo de los intervalos de confianza, por defecto se fija al 5%.
* **useMiss**: lógico, indica cuando usar o no los valores ausentes como categoría válida para el cálculo y reporte, valor por defecto FALSE, no usar valores ausentes como válidos.
* **verbose:** lógico, indica cuando se dejan los archivos intermedios generados en el proceso de estimación, y se permite la generación en el log de las notas de compilación

### USO

Para la estimación de promedios la macro a usar es ***%ESTIMA\_RAZON***. Para ilustrar el uso de la macro se presenta la invocación que se realiza para la estimación del promedio de dientes sanos; para esto la información tanto de las variables del diseño muestral como de los indicadores clínicos de los cinco cuestionarios debe estar consolidada en un data set de SAS, en este ejemplo se denomina DATOS en la librería work. De manera tal que la invocación de la macro para la estimación por grupos de edad en dentición temporal se realiza de la siguiente manera:

%***ESTIMA\_RAZON***(CampoCol = TIPO\_TEMPORAL, CampoFil = EDAD,

yk = IDS\_POR\_COP\_P, zk = PER\_EXAM, zkd = PER\_EXAM,

BaseIn = datos, BaseOut = creadas.IDS\_POR\_COP\_P\_EDAD2,

Estrato = st\_i st\_ii st\_iii st\_iv st\_v st\_vi,

UM = id\_i id\_ii id\_iii id\_iv id\_v id\_vi,

Diseno = dis\_i dis\_ii dis\_iii dis\_iv dis\_v dis\_vi,

N = n\_i n\_ii n\_iii n\_iv n\_v n\_vi,

nm = nm\_i nm\_ii nm\_iii nm\_iv nm\_v nm\_vi,

p = p\_i p\_ii p\_iii p\_iv p\_v p\_vi,

fkFinal = fex\_fin,

useMiss = FALSE, alphaSig = **0.05**, verbose = FALSE);

Debido a que las estimaciones para estos indicadores se realizan por tipos de dentición y para cada tipo de dentición se utilizan ciertas edades índice, se define una variable dummy que para este ejemplo es denominada TIPO\_TEMPORAL que indica los pacientes que están en las edades índice usadas en este caso para dentición temporal. Usando la macro de esta manera al final se realiza un filtro de los resultados para TIPO\_TEMPORAL = 1 para obtener los resultados en dentición temporal.

Los grupos de edad para reporte son generados en una variable categórica que para este ejemplo se denomina edad que tiene la codificación definida para el reporte. En la variable yk se usa el indicador clínico que recopila la cantidad de dientes sanos en la dentición permanente, en zk y zkd se usa una variable dummy que indica si el paciente le fue aplicado el examen clínico.

Las variables que definen el diseño muestral y que son usadas en los parámetros UM, DISENO, N, nm, p y fkFinal se encuentran definidas en la documentación de las tablas de datos.

### RESULTADOS

En la Tabla 1 se presenta un ejemplo de la salida de la invocación de la macro presentada en la sección anterior. Los primeros campos de la tabla corresponden a las variables usadas en los parámetros CampoCol y CampoFil, para el ejemplo esto corresponde a TIPO\_TEMPORAL y EDAD.

Tabla 1: Ejemplo de salida de la macro para estimación de razones

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **TIPO\_**  **TEMPORAL** | **EDAD** | **N\_R** | **SUMA\_PESOS** | **P** | **P\_EE** | **CVE\_P** | **LI\_P** | **LS\_P** | **R\_ESTIMACION** | **R\_EE** | **CVE\_R** | **LI\_R** | **LS\_R** |
| 1 | 4 | 2183 | 861113.94 | 0.02701 | 0.003070 | 0.11364 | 0.02100 | 0.03303 | 0.99978 | 0.000166 | 0.00017 | 0.99946 | 100.011 |
| 1 | 5 | 2046 | 873245.12 | 0.02739 | 0.002760 | 0.10074 | 0.02199 | 0.03280 | 0.99986 | 0.000097 | 0.00010 | 0.99967 | 100.005 |
| 1 | 6 | 1809 | 872072.40 | 0.02736 | 0.002197 | 0.08030 | 0.02305 | 0.03166 | 0.99935 | 0.000302 | 0.00030 | 0.99876 | 0.99994 |
| 1 | 7 | 2855 | 11460569.00 | 0.35952 | 0.004956 | 0.01378 | 0.34981 | 0.36924 | 0.99891 | 0.000081 | 0.00008 | 0.99876 | 0.99907 |
| 1 | 8 | 1686 | 5937327.47 | 0.18626 | 0.002731 | 0.01466 | 0.18090 | 0.19161 | 0.98700 | 0.001977 | 0.00200 | 0.98313 | 0.99087 |
| 1 | 9 | 2291 | 9121395.41 | 0.28614 | 0.004556 | 0.01592 | 0.27721 | 0.29507 | 0.90159 | 0.003284 | 0.00364 | 0.89516 | 0.90803 |
| 1 | 10 | 1180 | 2751438.70 | 0.08631 | 0.001897 | 0.02198 | 0.08259 | 0.09003 | 0.59450 | 0.008482 | 0.01427 | 0.57787 | 0.61112 |
| 1 | T | 14050 | 31877162.04 | 100.000 | 0.000000 | 0.00000 | 100.000 | 100.000 | 0.93400 | 0.001801 | 0.00193 | 0.93047 | 0.93753 |

Para una mejor interpretación de los resultados entregados en la tabla, se describen los campos que la componen a continuación:

* **N\_R: corresponde** a total de pacientes encuestados para cada una de las combinaciones de categorías de CampoCol y CampoFil, teniendo en cuenta el dominio generado por la variable definida en zkd.
* **SUMA\_PESOS:** Corresponde a la estimación del total de población a nivel nacional una de las combinaciones de categorías de CampoCol y CampoFil, teniendo en cuenta el dominio generado por la variable definida en zkd y el factor de expansión final usado para las estimaciones.
* **P:** Corresponde a la estimación de la proporción de población en una de las combinaciones de categorías de CampoCol y CampoFil que denotaremos , teniendo en cuenta el dominio generado por la variable definida en zkd.
* **P\_EE:** Corresponde a la estimación del error estándar de la estimación de la proporción que denotaremos , y es una de las medidas de calidad de la estimación que se reportan para los diferentes indicadores.
* **CVE\_P:** Corresponde al coeficiente de variación estimado para la proporción , se define como y es otra medida de calidad de la estimación.
* **LI\_P:** Es el límite inferior del intervalo de confianza para la proporción al nivel (1 - alphaSig) definido en la invocación de la macro, por defecto el valor usado para la estimación de alphaSig es de 0.05 que genera intervalos al 95% de confianza. El límite inferior es definido como donde corresponde al percentil de la distribución normal estándar.
* **LS\_P:** Es el límite superior del intervalo de confianza para la proporción definido como
* **R\_ESTIMACION:** Corresponde a la estimación de la razón para cada una de las combinaciones de categorías de CampoCol y CampoFil, esta razón corresponde al cociente de los totales estimados con las variables yk y zk definidas en la invocación de la macro, se denotara por .
* **R\_EE:** Corresponde a la estimación del error estándar de la estimación de la razón que denotaremos , y es una de las medidas de calidad de la estimación que se reportan para los diferentes indicadores.
* **CVE\_R:** Corresponde al coeficiente de variación estimado para la razón , se define como y es otra medida de calidad de la estimación.
* **LI\_R:** Es el límite inferior del intervalo de confianza para la razón definido como .
* **LS\_R:** Es el límite superior del intervalo de confianza para la razón definido como .