

## **APÉNDICE A**

### **DISEÑO Y COBERTURA DE LA MUESTRA**



---

## APÉNDICE A

### DISEÑO Y COBERTURA DE LA MUESTRA

#### A.1 Introducción

Para la Encuesta Nacional de Demografía y Salud 2000 (ENDS 2000) se utilizó una submuestra de la **Muestra Maestra Nacional** de PROFAMILIA, conformada por 1,023 segmentos con 10 viviendas en promedio en cada segmento. Los segmentos están distribuidos proporcionalmente en las zonas urbana y rural (cabecera y resto) de 61 municipios de todos los tamaños y regiones del país.

El universo de la muestra estudiada representa aproximadamente el 97 por ciento de la población nacional de las mismas características. No incluye los antiguos territorios nacionales (áreas de gran extensión y poca población), por los costos implicados en el estudio de una muy pequeña fracción de la población del país.

El tamaño de la muestra, alrededor de 11,000 hogares, permite hacer estimaciones con un error estándar relativo inferior a 0.1, de todas las variables estudiadas en las mujeres en edad fértil (MEF entre 15 y 49 años) a nivel nacional, de aquellas características con una probabilidad de ocurrencia de 0.07 ó más a nivel de las cinco grandes regiones, y de las características de las MEF con probabilidad de 0.19 ó más a nivel subregional (véase el Cuadro A.1).

Para los niños nacidos después del primero de enero de 1995, la muestra permite estimaciones con un error relativo menor o igual a 0.1 de las características que a nivel nacional tengan una probabilidad de ocurrencia de 0.03 ó más, y que a nivel regional la probabilidad sea de 0.16 ó más.

#### A.2 La Muestra Maestra

La **Muestra Maestra** de Profamilia está conformada por seis submuestras nacionales de 1,023 segmentos y aproximadamente 11,000 hogares cada una, distribuidos en 61 municipios seleccionados probabilísticamente del total de municipios de los departamentos del país (definición de 1990). Estas características de la muestra están asociadas con la precisión deseada, con consideraciones logísticas, operativas, de costo de su preparación y utilización, y con el uso esperado de la misma.

Para la selección de cada una de las muestras y presentación rutinaria de los resultados de las investigaciones, se dividió el país (departamentos de 1990) en 6 regiones con un mínimo de 4 millones de habitantes y en 14 subregiones con un mínimo de 800 mil habitantes. Los resultados de investigaciones como las encuestas de demografía y salud son presentados, para efectos de comparación, utilizando el criterio tradicional de 5 regiones, sin afectar la representatividad de la muestra maestra, pues la región adicional (Antioquia) aparece como una subdivisión de la Región Central.

La característica de "**maestra**" tiene que ver con la posibilidad de ser utilizada total o parcialmente, en forma simultánea, sucesiva o acumulativa, para múltiples investigaciones, independientemente de sus objetivos específicos.

## A.3 Diseño y Selección

### Tamaño y Precisión

El Cuadro A.1. presenta, para una submuestra de la **Muestra Maestra**, las relaciones entre tamaño de universo (país, región, subregión), tamaño de muestra y probabilidad de ocurrencia de un fenómeno cualquiera, con un error estándar relativo del 10 por ciento (ESrel = 0.1).

Cuadro A.1 Relaciones en la muestra maestra de Profamilia		
Relaciones <i>a priori</i> entre tamaño del universo, tamaño de la muestra y probabilidad de ocurrencia de un fenómeno para un mismo nivel de precisión (error estándar relativo de 0.1)		
Universo de estudio	Probabilidad de ocurrencia <sup>1</sup>	
	Mujeres (12,531)	Niños (4,670)
<b>A nivel del país</b>		
Población total	0.0122	0.0304
Subgrupos de 50 por ciento	0.0242	0.0590
Subgrupos de 30 por ciento	0.0397	0.0946
Subgrupos de 20 por ciento	0.0584	0.1355
<b>A nivel de región</b>		
Población total	0.0692	0.1583
Subgrupos de 50 por ciento	0.1295	0.2730
Subgrupos de 30 por ciento	0.1987	0.3854
Subgrupos de 20 por ciento	0.2710	0.4852
<b>A nivel de subregión</b>		
Población total	0.1879	0.3688
Subgrupos de 50 por ciento	0.3162	0.5388
Subgrupos de 30 por ciento	0.4354	0.6598
Subgrupos de 20 por ciento	0.5359	0.7461

Nota: Los cálculos fueron realizados con base en fórmulas deducidas para muestreo de conglomerados desiguales, con un error estándar relativo de 0.1 y suponiendo un efecto de los conglomerados en el diseño (EDIS en español; deff en Inglés) de 1.5 a nivel nacional y regional y de 2 a nivel subregional.

<sup>1</sup>En una submuestra de 12,659 hogares

### Unidades de Muestreo

Las Unidades Primarias de Muestreo (UPM) son los municipios de 7,000 o más habitantes de los departamentos del país. Los antiguos territorios nacionales, convertidos en 1993 en departamentos, no forman parte del universo de esta muestra. Los municipios de menor tamaño se unieron con uno contiguo del mismo departamento. El tamaño mínimo de población de la UPM se determinó con base en la cantidad de segmentos necesarios para que un municipio de la **Muestra Maestra** represente el estrato al cual pertenece. Las Unidades Secundarias del Muestreo (USM) son los segmentos, o grupos contiguos de viviendas, con un tamaño promedio de 10 en una submuestra y de 60 en el conjunto de la **Muestra Maestra**.

Un total de 11 municipios (UPM) fueron de certeza, es decir, tuvieron probabilidad 1 de selección, por su tamaño similar o superior al tamaño promedio de los estratos. Estas unidades son los municipios más grandes del país.

## Estratificación

Para una selección óptima (menor error de muestreo) de las UPM de no certeza, se conformaron 50 estratos, agrupando a nivel de subregión las unidades más parecidas en sus características. Los criterios de estratificación fueron la población total, el tamaño de la población de la cabecera municipal (nivel de urbanización), el porcentaje de población rural, el departamento y características socioeconómicas generales.

## Selección de UPM de No Certeza

Se utilizó la técnica de "**selección controlada**" para la obtención probabilística de una UPM en cada estrato<sup>1</sup>. Esta técnica permite el incremento de la precisión de la muestra mediante el uso de controles o variables de estratificación adicionales. Dichos controles, aplicados a nivel de región, fueron el departamento al cual pertenecían las UPM y un índice de desarrollo social conformado mediante la combinación de variables sobre escolaridad de la población, disponibilidad de recursos de salud (camas hospitalarias, horas médicas y personal de enfermería), y disponibilidad de energía eléctrica.

## Selección de Segmentos en cada UPM Seleccionada

Utilizando como marco de muestreo los mapas y listados del Censo Nacional de Población de 1985 disponibles en el Departamento Administrativo de Estadística (DANE), se procedió de la siguiente manera:

- Se estimó el número de áreas, de 60 viviendas en promedio, que tendría la muestra en cada UPM seleccionada. Tal asignación se hizo en función del tamaño de la población del estrato respectivo en el caso de las UPM de no certeza, y de la población de las UPM de certeza. El número de segmentos urbanos y rurales (cabecera y resto) en cada UPM se definió con base en la composición urbano-rural de la población general de los estratos y de las UPM de certeza.

Las 60 viviendas de un área corresponden a los 6 segmentos, de 10 viviendas en promedio, de las respectivas 6 submuestras de la **Muestra Maestra**.

- Se seleccionó independientemente para cada UPM la muestra de áreas y éstas se dividieron en los segmentos de las varias submuestras.

Inicialmente, en cada UPM seleccionada, se listaron por área mínima de empadronamiento del censo de población de 1985 (sector, sección o manzana) las viviendas censadas, se calculó el número de las áreas teóricas de 60 viviendas en promedio, se elaboró una frecuencia acumulada de áreas teóricas separada por cabecera y resto, se calculó un intervalo de muestreo (áreas existentes sobre áreas a ser seleccionadas), y se seleccionaron las áreas de la muestra aplicando sistemáticamente el intervalo de muestreo a partir de un arranque aleatorio.

Se identificaron en la cartografía, los sectores, secciones y manzanas que contenían las áreas de la muestra que posteriormente fueron visitados para levantar el croquis con las viviendas y hogares existentes; se construyeron, sobre los croquis, segmentos con un promedio de 10 viviendas utilizando todo tipo de límites, divisiones o señales existentes en el terreno (calles y puntos de referencia en las zonas urbanas, y ríos, quebradas, carreteras, caminos, líneas férreas, montes, etc., en las zonas rurales). Finalmente, se distribuyeron aleatoriamente en las submuestras los segmentos construidos en cada área de muestreo.

---

<sup>1</sup>Goodman, R., y Kish, L. "Controlled Selection, A Technique in Probability Sampling", JASA, 1950.

## Probabilidades de Selección

El diseño de la muestra está orientado para generar la misma probabilidad final de selección a nivel de cada persona, hogar y segmento que la conforma. Tales probabilidades son:

### Primera etapa $f_a$ :

$$\frac{\text{población de la unidad primaria}}{\text{población del estrato}}$$

### Segunda etapa $f_b$ :

$$\frac{\text{segmentos seleccionados en unión primaria}}{\text{segmentos existentes en unidad primaria}}$$

### Probabilidad Final:

$$f = f_a \times f_b$$

Esta probabilidad final es igual porque el número de segmentos seleccionados en la unidad primaria es proporcional al tamaño del estrato de selección (al cual pertenecen) dentro del universo total. En las unidades de certeza, la probabilidad de la primera etapa es 1, y la probabilidad final también es igual, por la citada proporcionalidad en el número de segmentos seleccionados.

La igual probabilidad permite, en un principio, que los valores muestrales sean estimaciones no sesgadas para el universo del estudio. No obstante, algunas imperfecciones en el marco de muestreo construido a partir de los materiales del Censo Nacional de Población de 1985, problemas de no cobertura o de no respuesta en las investigaciones y los ajustes de los totales muestrales por zona urbana-rural a proyecciones de la población a la fecha de las encuestas, dan lugar, usualmente, a variaciones de cierta importancia en las probabilidades finales de selección, que obligan a aplicar ponderaciones a los valores muestrales.

## Composición de la Muestra

La submuestra de la **Muestra Maestra** utilizada en este estudio tiene la composición que se presenta en el Cuadro A.2, en términos de unidades primarias de muestreo y número de segmentos, por departamentos, subregion y región del país.

## Actualización de la Muestra

La **Muestra Maestra** de PROFAMILIA contempla diferentes estrategias para garantizar una actualización aproximada con respecto a los mayores cambios poblacionales que se van registrando a través del tiempo. En su diseño de áreas, la muestra dió probabilidad efectiva de selección a la totalidad de la superficie de su universo de investigación (antiguos departamentos), sin dejar por fuera de esta asignación probalística ninguna área urbana o rural, independientemente de su volumen de población (cero población en algunos casos). Para este efecto, todas las áreas despobladas, según el censo de 1985, fueron **anexadas** al área vecina más cercana (límitrofe) que tuviera alguna población. De esta manera se garantizó que los mayores cambios registrados, tanto en las áreas pobladas como en las áreas inicialmente despobladas, se captaran aproximadamente a través de los segmentos de la muestra

**Cuadro A.2 Una submuestra de la muestra maestra de Profamilia**

Número de departamentos, unidades primarias de muestreo (UPM) y segmentos por región y subregión

Región y subregión	Departamentos	Unidades primarias de muestreo	Número de segmentos		
			Urbanos	Rurales	Total
<b>Región Atlántica</b>					
Guajira-Cesar-Magdalena	3	5	35	31	66
Atlántico-Norte de Bolívar	3	5	71	8	79
Bolívar (resto)-Sucre-Córdoba	3	6	32	38	70
<b>Región Oriental</b>					
Santander-Norte de Santander	2	6	48	31	79
Boyacá-Cundinamarca-Meta	3	8	42	58	100
<b>Región Central</b>					
Antioquia sin Medellín	1	5	19	36	55
Medellín (Área Metropolitana)	1	1	68	4	72
Caldas-Risaralda-Quindío	3	5	45	21	66
Tolima-Huila-Caquetá	2	6	37	34	71
<b>Región Pacífica</b>					
Cauca-Nariño	2	4	21	32	53
Valle del Cauca	1	4	29	15	44
Litoral Pacífico	3	4	13	13	26
Cali (con Yumbo)	1	1	57	1	58
<b>Bogotá ( con Soacha)</b>	1	1	183	1	183
<b>Total</b>	<b>25</b>	<b>61</b>	<b>700</b>	<b>323</b>	<b>1,023</b>

En el caso particular de la submuestra utilizada para la **ENDS 2000**, en octubre de 1999 se visitaron los segmentos de las 4 UPM con mayor tamaño poblacional y se actualizó tanto la cartografía como la identificación de las viviendas y hogares de estos segmentos. En esta misma oportunidad se actualizó la muestra en los segmentos del municipio de Armenia, uno de los más afectados por el terremoto que azotó el eje cafetero en enero de 1999 y, atendiendo los graves problemas de orden público que afectan el sur de Bolívar, el municipio de Majagual fue remplazado por otro municipio de la muestra perteneciente al mismo estrato de selección.

En total, previo al trabajo de campo se visitaron 373 segmentos de un total de 1,023, es decir el 37 por ciento de la muestra, lo cual contribuyó a agilizar la ubicación de los segmentos y la identificación de viviendas y hogares a ser entrevistados en los centros urbanos más susceptibles a presentar variaciones en su desarrollo urbano. Cuando se encontraron grandes crecimientos poblacionales en los segmentos de la muestra fue necesario submuestrear los nuevos hogares captados, generando por consiguiente variación en las probabilidades inicialmente asignadas.

Adicionalmente, con motivo de cada nueva encuesta, se introduce un ajuste en las probabilidades de selección, para tener en cuenta los cambios "proyectados" en la distribución de la población por zonas urbana y rural a nivel de estrato, a fin de captar con alguna aproximación la influencia, en las variables dependientes estudiadas, de los cambios en la estructura de la población por nivel de urbanización.

Por último, la actualización más importante se puede realizar con los datos y materiales de cada nuevo censo nacional de población aplicando la metodología sugerida por Kish, la cual permite realizar este proceso probabilísticamente, pero reteniendo una buena proporción de las unidades primarias de muestreo seleccionadas en un comienzo, a fin de poder seguir usando los listados (actualizados) y materiales cartográficos de estas unidades<sup>2</sup>.

<sup>2</sup>Kish, L. **Survey Sampling**, 2<sup>nd</sup> edition, John Wiley and Sons, New York, 1965

## A.4 Procedimientos de Estimación

Se denomina Procedimientos de Estimación a la metodología de tratamiento de los datos de una encuesta por muestreo para garantizar estimaciones insesgadas del universo investigado. Son fuentes de sesgo las diferentes probabilidades de selección y la no cobertura diferencial de la muestra seleccionada entre grupos de población de distinta condición según estrato socioeconómico.

Las variables probabilidades de selección se controlan ponderando los valores muestrales por el recíproco de la probabilidad final de cada persona encuestada, o por un valor equivalente. Dicho recíproco, es el “Factor Básico de Expansión”, el cual restituye aproximadamente los valores absolutos del universo estudiado. El factor equivalente, puede elaborarse alrededor de la unidad, para mantener los resultados del estudio en valores muestrales (sin expansión), pero corrigiendo el efecto de las diferentes probabilidades .

La muestra de hogares de Profamilia tiene, por diseño, igual probabilidad final de selección, y un factor básico único de expansión, o de ajuste alrededor de la unidad. Sin embargo, la diferente tasa de cobertura de la muestra por estrato socioeconómico y por zona, origina probabilidades finales desiguales de la muestra encuestada. Un factor de recuperación de los hogares y establecimientos no encuestados, a nivel de segmento, los primeros, y por municipio y zona, los segundos, cuando es presumible que la muestra encuestada es similar, en promedio, a la no encuestada, corrige el eventual sesgo de la no cobertura diferencial, si se mantiene el objetivo de inferir los resultados al universo de estudio original.

Adicionalmente, es posible mejorar la calidad de las estimaciones, corrigiendo la estructura por zona de residencia, de la muestra encuestada, en función de la estructura proyectada del censo de población a la fecha de la encuesta, la cual se considera más real y exacta, debido a la relación que existe entre estas variables y los indicadores estudiados:

- (1) **Probabilidad final de la muestra básica:**

$$f = f_a \times f_b$$

- (2) **Factor básico de expansión (Recíproco de la probabilidad final):**

$$F = \frac{1}{f}$$

- (3) **Ajuste por no cobertura de hogares:**

$$H = \frac{\text{hogares seleccionados en el segmento}}{\text{hogares encuestados en el segmento}}$$

- (4) **Ajuste por no cobertura de MEF:**

$$M = \frac{\text{MEF existentes en hogares encuestados en el segmento}}{\text{MEF encuestadas en hogares encuestados en el segmento}}$$

(5) **Ajuste por estructura urbano/rural:**

-Factor urbano U:

$$\frac{\text{porcentaje de población urbana proyectada en el estrato}}{\text{porcentaje de población urbana de la muestra en el estrato ajustada por no cobertura}}$$

-Factor rural R:

$$\frac{\text{porcentaje de población rural proyectada en el estrato}}{\text{porcentaje de población rural de la muestra en el estrato ajustada por no cobertura}}$$

(6) **Estandarización del factor de expansión F' :**

$$\frac{F \text{ (de las personas y hogares en un segmento)}}{F \text{ (promedio para el país)}}$$

(7) **Factor Final:** Factor básico ajustado por no cobertura, por estructura urbano - rural y estandarizado.

$$F_f = F \times H \times M \times U \times F'$$

**Estimaciones de Razón:**

Son los resultados del cociente de dos variables (promedios, proporciones, tasas), que en este diseño particular tienen el carácter de variables aleatorias. La razón se expresa:

$$r = \frac{y}{x} = \frac{\sum_h \sum_\alpha \sum_\beta \sum_\gamma \sum_\delta Y_{h\alpha\beta\gamma\delta} W_{h\alpha\beta}}{\sum_h \sum_\alpha \sum_\beta \sum_\gamma \sum_\delta X_{h\alpha\beta\gamma\delta} W_{h\alpha\beta}}$$

En donde:

- h = 1,2,3, ..., H son los diferentes estratos
- $\alpha = 1,2,3, \dots, a$  son las unidades primarias de la muestra en el h-ésimo estrato
- $\beta = 1,2,3, \dots,$  son los segmentos seleccionados en la  $\alpha$ -ésima unidad primaria del h-ésimo estrato
- $\gamma = 1,2,3, \dots,$  son los hogares del  $\beta$ -ésimo segmento, de la  $\alpha$ -ésima unidad primaria, del h-ésimo estrato
- $\delta = 1,2,3, \dots,$  son las personas del hogar  $\gamma$ -ésimo, del  $\beta$ -ésimo segmento, de la  $\alpha$ -ésima unidad primaria, del h-ésimo estrato

$Y_{h\alpha\beta\gamma\delta}$ 

es el valor de la variable de estudio en cada persona de la muestra.

Si la variable es continua (peso, edad, estatura, etc), es el valor respectivo en cada persona; si la variable es binomial (o convertida en binomial), cada persona que tiene la característica estudiada vale 1.

 $X_{h\alpha\beta\gamma\delta}$ 

es cada persona de la muestra

 $W_{h\alpha\beta}$ 

Factor final

## A.5 Rendimiento de la Muestra

En el Cuadro A.3 se presentan las tasas de respuesta por región y por zona de residencia. En el 7.3 por ciento de las 12,659 viviendas visitadas no se encontraron hogares porque la vivienda estaba desocupada, cambió de uso o fue demolida. En total se encontraron 11,736 hogares y en 10,907 se logró completar la entrevista, para una tasa de respuesta del 93 por ciento al excluir del cálculo los hogares no disponibles para entrevistas por las razones enunciadas. De un total de 12,531 mujeres elegibles se logró entrevistar 11,585, obteniéndose una tasa de respuesta de 93 por ciento.

La tasa de respuesta combinada de hogares y mujeres fue del 86 por ciento por ciento, bastante satisfactoria para este tipo de encuesta, sobre todo tomando en cuenta las difíciles condiciones que presentan algunos departamentos de Colombia para el trabajo de campo. Se mejoró substancialmente la tasa de respuesta en Bogotá tanto para los hogares como para las entrevistas individuales mientras que en las otras regiones se mantuvo la cobertura de 1995, aunque en la Región Pacífica la respuesta de las mujeres fue un poco menor.

Cuadro A.3. Tasas de respuesta en la ENDS 2000 por región y zona de residencia								
Tasas de respuesta en la encuesta de hogares y en la individual por región y zona de residencia, Colombia 2000								
	Zona		Región					Total
	Urbana	Rural	Atlántica	Oriental	Central	Pacífica	Bogotá	
<b>Hogares Potenciales</b>								
<i>Entrevista Completa</i>	86.8	84.4	90.5	87.5	85.2	85.5	80.7	86.2
Rechazo	2.9.0	0.2	0.5	1.7	2.2	2.3	5.1	2.2
Ausente momentáneo	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Ausente temporal	0.1	0.1	0.0	0.1	0.1	0.1	0.1	0.1
Ausente no recuperable	4.8	3.0	2.4	2.7	4.2	5.1	8.1	4.3
Vivienda desocupada	5.3	11.9	6.5	7.9	8.2	7.0	5.9	7.2
Cambio de uso	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0
Otras razones	0.0	0.3	0.1	0.1	0.1	0.1	0.0	0.1
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Número de hogares	9,026	3,633	2,891	2,202	3,435	2,163	1,968	12,659
<b>Mujeres Elegibles</b>								
<i>Entrevista Completa</i>	92.3	93.0	94.0	92.6	92.7	91.7	89.8	92.5
Rechazo	1.3.0	0.3	0.4	1.8	0.5	1.3	2.5	1.1
Ausente momentáneo	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0
Ausente temporal	0.1	0.0	0.1	0.0	0.2	0.0	0.1	0.1
Ausente no recuperable	5.5	4.8	4.7	4.4	5.3	5.9	6.9	5.3
Entrevista Incompleta	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.0
Otra razón	0.7	1.7	0.8	1.0	1.2	1.0	0.7	0.9
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Número de mujeres	9,536	2,995	3,296	2,049	3,254	2,088	1,844	12,531
<b>Tasa de respuesta</b>								
Hogares	91.7	96.2	96.9	95.1	92.9	92.0	85.8	92.9
Mujeres	92.3	93.0	94.0	92.6	92.7	91.7	89.8	92.5
Total	84.6	89.5	91.1	88.1	86.1	84.4	77.0	85.9

## A.6 Submuestreo para el Componente de Salud Visual

Con el objeto de conocer el estado de salud visual de la población colombiana, se realizó un examen de ojos a una muestra compuesta por dos grandes grupos poblacionales:

- Niños no escolarizados de 6 a 11 años y personas de 50 años y más que fueron captados en los hogares de la muestra utilizada para la **ENDS 2000**. Para estos dos grupos la probabilidad final y factor básico de expansión son los mismos de la muestra general de hogares.
- Niños de 6 a 11 años que en el momento de la encuesta estuvieran cursando cualquier año de educación primaria, cuya salud visual fue evaluada clínicamente en los establecimientos en los cuales estudian.

Para la ubicación de los escolares se dispuso de una lista de establecimientos educativos previamente seleccionados probabilísticamente entre los existentes en los municipios de la muestra nacional de la **ENDS 2000**. Para cada establecimiento el(la) optómetra solicitó la lista de los estudiantes de todos los grados de primaria (de primero a quinto, por grupo). En forma separada para cada grado, enumeró consecutiva y en forma continua los estudiantes de todos los grupos, determinó el número de estudiantes que debían ser seleccionados en cada grado y construyendo un intervalo de muestreo con arranque aleatorio seleccionó sistemáticamente los estudiantes a ser examinados. Finalmente verificó la edad de los niños seleccionados controlando que estuviera comprendida entre los 6 y los 11 años. Los niños que no correspondieron a este rango de edad fueron descartados y en su remplazo se tomó al niño siguiente del listado.

## A.7 Procedimientos de Estimación para el Componente de Salud Visual

A continuación se presentan las probabilidades de selección de la muestra de establecimientos educativos dentro de cada unidad primaria (UPM) de la muestra básica de Profamilia, y de los niños encuestados dentro de cada establecimiento seleccionado. Se define igualmente el factor básico de expansión de los resultados para corregir el efecto de las probabilidades finales desiguales, y expandir, si se desea, los resultados absolutos del estudio de salud visual. También se presentan los factores para la recuperación de los establecimientos seleccionados y no encuestados; se determinan los factores de ajuste de la estructura por edad y sexo, y por zona y región y se indica la forma de calcular el factor final de ponderación de la muestra alrededor de la unidad.

- La probabilidad básica de selección y el factor básico de expansión son los mismos de la muestra general.

$$p = \frac{1}{f} = \frac{1}{f_a \times f_b}$$

- La probabilidad de selección de cada establecimiento, dentro de la zona del municipio en que está ubicado, vale decir, zona urbana, zona rural concentrada (Centro Poblado) o zona rural dispersa, es proporcional al número de grupos de estudiantes:

$$p_{E_1} = \frac{\text{Grupos en establecimiento seleccionado}}{\text{Grupos en el mismo municipio y zona}} \times \text{Establecimientos seleccionados en mismo municipio y zona}$$

- La probabilidad de selección de estudiantes dentro de cada establecimiento seleccionado, por grado:

$$p_{E_2} = \frac{\text{Estudiantes valorados en un grado}}{\text{Estudiantes existentes en dicho grado}}$$

(Esta probabilidad calculada para cada grado, recupera los estudiantes seleccionados -existentes- y no examinados).

- Probabilidad de selección de los niños escolarizados, en el municipio:

$$P_E = P_{E_1} \times P_{E_2}$$

- Factor de expansión al municipio (recíproco de la probabilidad de selección de los niños en el municipio):

$$R = \frac{1}{P_E}$$

- Recuperación de establecimientos seleccionados (existentes) y no examinados, por zona, por municipio.

$$F_A = \frac{\text{Total de estudiantes de establecimientos seleccionados}}{\text{grupos de estudiantes examinados en el mismo municipio y zona}}$$

- Factor de Expansión al municipio, ajustado por no encuesta de establecimientos

$$R_A = R \times F_A$$

- Factor de ajuste de la estructura por edad y sexo de la muestra ponderada, escolarizada y no escolarizada.

$$F_{ES} = \frac{\text{Proyección población universo (proyección censal) en grupo de edad y sexo}}{\text{Proporción muestra ponderada en mismo grupo de edad y sexo}}$$

- Factor de ajuste de la estructura por zonas urbana y rural de la muestra ponderada, escolarizada y no escolarizada, ajustada por edad y sexo.

$$F_{UR} = \frac{\text{Proporción de población del universo (proyección) en cada zona}}{\text{Proporción de la muestra ponderada, ajustada por edad y sexo en la misma zona}}$$

- Reducción de factores de expansión a valores alrededor de la unidad

$$F' = \frac{\text{Factor original de expansión}}{\text{Factor de expansión promedio}}$$

$$\text{Factor de expansión promedio} = \frac{\text{Universo poblacional}}{\text{Muestra poblacional seleccionada}}$$

- Ponderación final de la muestra (con factores alrededor de la unidad), a nivel de un grupo de edad y sexo, de una determinada zona, de un estrato.

#### *Muestra escolarizada*

$$R_F = R_A \times F_{ES} \times F_{UR} \times F'$$

#### *Muestra no escolarizada*

Ponderación de la muestra no escolarizada por Factor de Ajuste de estructura edad-sexo

$$R_F = F_f \times F_{ES} \times F_{ur} \times F'$$