



INTRODUCCIÓN

En este documento se describe el diseño de muestreo utilizado en la Encuesta Nacional de Tutoría en Educación Básica y Educación Media Superior. Se describen los procedimientos de construcción y depuración de los marcos muestrales; se presentan la confiabilidad y precisión esperadas de los estimadores, de acuerdo a los tamaños de muestra establecidos; se explican los cálculos de los factores de expansión y los ajustes aplicados por la corrección de la cobertura de los marcos muestrales, y sesgos por selección y por no respuesta. Los errores de muestreo fueron calculados a partir del enfoque uni-etápico.



Tabla de contenido

I. POBLACIÓN OBJETIVO, COBERTURA Y DISEÑO	3
1.1 Población objetivo	3
1.2 Cobertura geográfica.....	3
1.3 Diseño de la muestra.....	3
II. MARCO DE LA ENCUESTA.....	4
2.1 Formación de las Unidades Primarias de Muestreo (UPM).....	4
2.2 Estratificación	4
III. ESQUEMA DE MUESTREO	5
IV. TAMAÑO DE LA MUESTRA	6
V. ASIGNACIÓN DE LA MUESTRA	7
VI. SELECCIÓN DE LA MUESTRA	8
6.1 Tutorados	8
6.2 Tutores	8
VII. SELECCIÓN DE LA MUESTRA	10
7.1 Tutorados	10
7.2 Tutores	10
7.3 Tasa de respuesta observada	10
7.4 Factores de expansión normalizados	10
VIII. ESTIMADORES	12
IX. ESTIMACIÓN DE LAS PRECISIONES	13
ANEXO A. PROCESO DE CONSTRUCCIÓN Y DEPURACIÓN DE LOS MARCOS MUESTRALES	15

I. POBLACIÓN OBJETIVO, COBERTURA Y DISEÑO

1.1 Población objetivo

Docentes de Educación Básica (EB) y Educación Media Superior (EMS) que ingresaron al servicio en el ciclo escolar 2014-2015, y personal de Educación Básica y Educación Media Superior que desempeñó la función tutora durante los ciclos 2014-2015 y 2015-2016.

1.2 Cobertura geográfica

La encuesta se diseñó para obtener estimaciones para tutorados de EB a nivel Nacional (sin incluir a Oaxaca y Chiapas) y 30 entidades federativas; para tutorados de EMS a nivel Nacional (sin incluir a Oaxaca y Chiapas) y 30 entidades federativas; para tutores de EB a nivel Nacional (sin incluir Oaxaca, Chiapas, Baja California Sur, Michoacán y Sonora) y 27 entidades federativas, y para tutores de EMS a nivel Nacional (sin incluir a Oaxaca, Chiapas y Guanajuato) y 29 entidades federativas.

1.3 Diseño de la muestra

El diseño de la encuesta se planteó probabilístico para obtener una muestra representativa de la población; se estableció para ser estratificado uni-étapico y las unidades de observación son los tutorados (EB y EMS) y tutores (EB y EMS).

II. MARCO DE LA ENCUESTA

2.1 Formación de las Unidades Primarias de Muestreo (UPM)

El diseño de la muestra es uni- etápico dentro de los estratos, por lo tanto, las unidades primarias de muestreo son las unidades de observación: tutorados y tutores.

2.2 Estratificación

La población objetivo se estratificó de acuerdo al siguiente esquema:

Docente o personal	Modalidad	Nivel Geográfico	Nivel Educativo	Grupo Desempeño	Estratos finales
Tutorados	Básica	30 entidades (excepto Oaxaca y Chiapas)	5 categorías: educación especial y otro, preescolar, primaria, secundaria e indígena ^b	2 categorías: AB y CD	194 ^a
	Media Superior	30 entidades (excepto Oaxaca y Chiapas)	1 categoría: media superior.	2 categorías: AB y CDE	60
Tutores	Básica	27 entidades (excepto Oaxaca, Chiapas, Baja California Sur, Michoacán y Sonora)	-	-	27
	Media Superior	29 entidades (excepto Oaxaca, Chiapas y Guanajuato)	-	-	29

^a Algunos estratos definidos por nivel geográfico, nivel educativo y desempeño estaban vacíos.

^b Indígena se consideró como una categoría independiente del nivel educativo.

En el **Anexo A** se describe el proceso para la construcción y depuración de los marcos muestrales.

III. ESQUEMA DE MUESTREO

La Encuesta Nacional de Tutoría en Educación Básica y Educación Media Superior fue diseñada bajo un esquema de muestreo probabilístico, estratificado y uni-etápico:

- es probabilístico ya que las unidades de muestreo tienen una probabilidad conocida y mayor a cero de ser seleccionadas;
- es estratificado, las unidades de muestreo se clasifican de acuerdo a la modalidad de educación, nivel geográfico, nivel educativo y grupo de desempeño y
- es uni-etápico porque la unidad de observación (tutorado o tutor) se selecciona en una etapa dentro de cada estrato.

IV. TAMAÑO DE LA MUESTRA

El tamaño de la muestra tiene un error teórico asociado a las estimaciones de δ al 95% de confianza, asumiendo muestreo aleatorio simple y asumiendo una población suficientemente grande.

El tamaño de muestra está definido por la siguiente fórmula:

$$n = \left(\frac{1}{\frac{\delta^2}{z_{\alpha/2}^2 \left(\frac{N}{N-1} \right) P(1-P)} + \frac{1}{N}} \right) \times \frac{1}{R}$$

donde:

n = tamaño de muestra de unidades de observación.

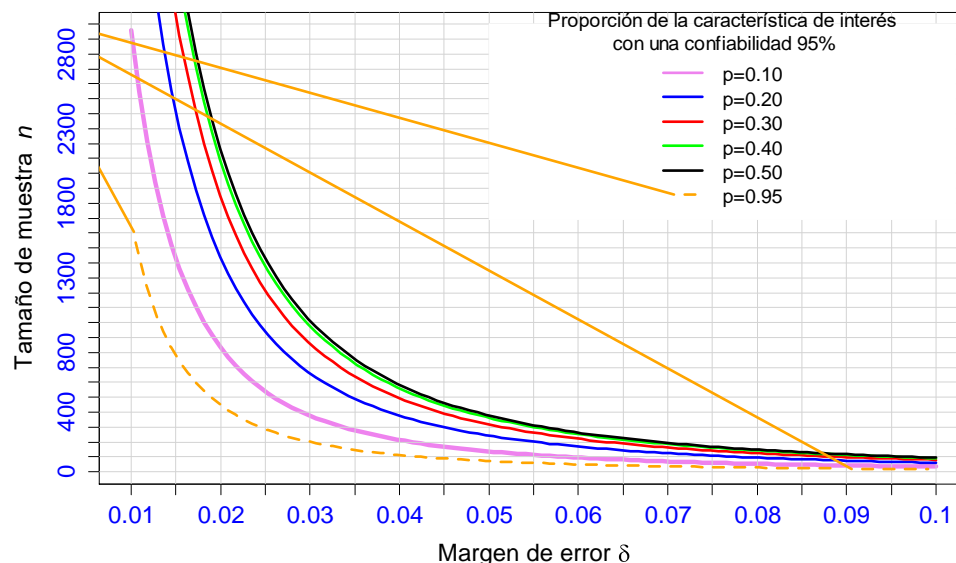
P = proporción a estimar.

N = tamaño de la población objetivo.

$z_{\alpha/2}$ = es el valor del cuantil de una distribución normal estándar que deja un área de $\alpha/2$ a la derecha del él, en particular $\alpha = 0.05$

δ = margen de error teórico al 95% de confianza bajo muestreo aleatorio simple. Semi-amplitud del intervalo de confianza.

R = tasa o porcentaje de respuesta esperada.



De acuerdo a la proporción de interés P a estimar (por ejemplo, la proporción de tutorados que consideran que su práctica docente ha mejorado a partir de las tutorías), la tasa de respuesta observada R , el tamaño N de la población objetivo y el tamaño de muestra n se tendrá un margen de error $\pm \delta$ del 95% de confianza, es decir, el margen de error estará determinado por la combinación de las cuatro cantidades mencionadas. Un porcentaje alto de respuesta garantiza lograr el margen de error δ establecido.

V. ASIGNACIÓN DE LA MUESTRA

La muestra básica fue de 2,000 entrevistas que se asignaron aproximadamente proporcional al tamaño de las cuatro poblaciones objetivo. Una vez determinado el tamaño de muestra en cada población objetivo, dentro de los estratos definidos se asignó proporcionalmente el número de entrevistas; también se consideró una sobremuestra de 2 entrevistas adicionales para cada uno de los estratos (en aquellos estratos donde la población fue suficiente). Este incremento de la muestra se debió que el operativo de campo consistió en citar a los tutorados o tutores en una sede en una fecha fija; esta restricción pudo no permitir alcanzar los tamaños de muestra definidos.

Docente o personal	Modalidad	Tamaño de muestra asignado (número de entrevistas)	Sobremuestra (porcentaje respecto al tamaño de muestra original)
Tutorados	EB	1500	381 (25.4%)
	EMS	500	120 (24.0%)
Tutores	EB	210	54 (25.7%)
	EMS	90	29 (32.0%)

VI. SELECCIÓN DE LA MUESTRA

La selección de la muestra para la Encuesta Nacional de Tutoría en Educación Básica y Educación Media Superior se realizó en forma independiente para cada modalidad y estrato, y separado para tutorados y para tutores:

6.1 Tutorados

De los N_{me} tutorados que integraron los marcos muestrales en ambas modalidades, se eligieron n_{me} tutorados con igual probabilidad dentro de cada estrato. La probabilidad de elegir al j —ésimo tutorado del e —ésimo estrato en la modalidad m es

$$P\{D_{mej}\} = \frac{n_{mej}}{N_{mej}}$$

El factor de expansión correspondiente se define como:

$$F_{mej} = \frac{1}{P\{D_{mej}\}} = \frac{N_{mej}}{n_{mej}}$$

donde:

N_{me} = número total de docentes del marco muestral, en el e —ésimo estrato (nivel y grupo de desempeño), en la modalidad m .

n_{me} = número de docentes seleccionados del marco muestral, en el e —ésimo estrato (nivel y grupo de desempeño), en la modalidad m .

$e = 1, 2, \dots, 194$ si $m =$ Educación Básica, y $e = 1, 2, \dots, 60$ si $m =$ Educación Media Superior.

6.2 Tutores

De las N_{me} tutores que integraron los marcos muestrales en ambas modalidades, se eligieron n_{me} tutores con igual probabilidad dentro de cada estrato. La probabilidad de elegir al j —ésimo tutor del e —ésimo estrato (entidad) en la modalidad m es

$$P\{T_{mej}\} = \frac{n_{mej}}{N_{mej}}$$

El factor de expansión correspondiente se define como:

$$F_{mej} = \frac{1}{P\{T_{mej}\}} = \frac{N_{mej}}{n_{mej}}$$

donde:

N_{me} = número total de tutores del marco muestral, en el e —ésimo estrato, en la modalidad m .



n_{me} = número de tutores seleccionados del marco muestral, en el e —ésimo estrato, en la modalidad m .

$e = 1, 2, \dots, 27$ si $m =$ Educación Básica, y $e = 1, 2, \dots, 29$ si $m =$ Educación Media Superior

VII. AJUSTE A LOS FACTORES DE EXPANSIÓN

Los factores de expansión antes descritos se modificaron debido a la no respuesta:

7.1 Tutorados

$$F_{mej}^* = \frac{N_{mej}}{n_{mej}^*}$$

donde:

N_{me} = número total de docentes del marco muestral, en el e – ésimo estrato, en la modalidad m .

n_{mej}^* = número de docentes efectivamente entrevistados, en el e – ésimo estrato (nivel y grupo de despeño), en la modalidad m .

$e = 1, 2, \dots, 194$ si $m =$ Educación Básica, y $e = 1, 2, \dots, 60$ si $m =$ Educación Media Superior. En algunos estratos podría no haberse logrado al menos una entrevista, por lo que los estratos sin entrevistas efectivas no contribuyen para la estimación de los totales correspondientes.

7.2 Tutores

Los factores de expansión correspondientes a los tutores no se calcularon debido las dificultades operativas en las sedes de entrevistas: las distancias considerables entre las sedes y los centros de trabajo de los tutores obstaculizaron la dispersión geográfica de la muestra. En particular, las estimaciones de los parámetros de interés de los tutores se consideraron como un estudio exploratorio sin representatividad de la población objetivo.

7.3 Tasa de respuesta observada

En la siguiente tabla se presentan las tasas de respuesta logradas en cada una de las cuatro poblaciones de estudio.

Docente o personal	Modalidad	Tamaño de muestra asignado (número de entrevistas)	Número de entrevistas logradas (porcentaje de respuesta respecto al tamaño de muestra original)
Tutorados	EB	1500	1314 (87.6%)
	EMS	500	429 (85.8%)
Tutores	EB	210	181 (86.2%)
	EMS	90	87 (96.7%)

7.4 Factores de expansión normalizados

La ponderación de los casos en la encuesta es utilizada para producir estimaciones. Utilizar factores de expansión para el análisis de las encuestas podría tener un efecto en las estimaciones puntuales y en los errores estándares de estas estimaciones si se utilizan técnicas estándar de Estadística.

Un enfoque que puede ayudar en el tratamiento de valores ponderados en una base de datos, es normalizar los factores de expansión. La normalización de los factores de expansión¹ implica multiplicar el factor de expansión de cada caso por una constante de modo que se sumen el tamaño de muestra n .

¹ Chambers, R. L. & Skinner, C. J. (Eds.). (2003). Analysis of survey data. John Wiley & Sons.

VIII. ESTIMADORES

Los estimadores de un total X de una característica de interés a nivel nacional de los tutorados se obtienen como

$$\hat{t} = \sum_{x_{ej} \in m} \sum_e \sum_j F_{mej}^* x_{mej} = \hat{t}_{EB} + \hat{t}_{EMS}$$

Los estimadores de un total X_{EB} de una característica de interés de los tutorados de EB se obtienen como

$$\hat{t}_{EB} = \sum_e \sum_j F_{ej}^* x_{ej}$$

Los estimadores de un total X_{EB} de una característica de interés de los tutorados de EMS se obtienen como

$$\hat{t}_{EMS} = \sum_e \sum_j F_{ej}^* x_{ej}$$

Los estimadores de un total X_{EB_e} de una característica de interés de los tutorados de EB en el estrato e —ésimo, $e = 1, 2, \dots, 194$, se obtienen como

$$\hat{t}_{EB_e} = \sum_j F_{ej}^* x_{ej}$$

Los estimadores de un total X_{EB_e} de una característica de interés de los tutorados de EMS en el estrato e —ésimo, $e = 1, 2, \dots, 29$, se obtienen como

$$\hat{t}_{EMS_e} = \sum_j F_{ej}^* x_{ej}$$

donde x_{mej} es la observación del j —ésimo tutorado del e —ésimo estrato en la modalidad m .

$e = 1, 2, \dots, 194$ si $m =$ Educación Básica, y $e = 1, 2, \dots, 60$ si $m =$ Educación Media Superior. En algunos estratos podría no haberse logrado al menos una entrevista, por lo que los estratos sin entrevistas efectivas no contribuyen para la estimación de los totales correspondientes.

Estos cálculos no se realizaron para los estimadores de los parámetros de estudio de los tutores.

IX. ESTIMACIÓN DE LAS PRECISIONES

Para la evaluación de los errores de muestreo de las principales estimaciones estatales y nacionales se usó el método del módulo de encuestas complejas de SPSS.

El estimador de la varianza del estimador de un total nacional de tutorados se calcula como

$$\hat{Var}(\hat{t}) = \sum_{x_{ej} \in m} \sum_e \left(1 - \frac{n_{me}^*}{N_{me}}\right) N_{me}^2 \frac{s_{me}^2}{n_{me}^*}$$

donde

$$s_{me}^2 = \sum_j \frac{(x_{mej} - \bar{x}_{me})^2}{n_{me}^* - 1}$$

x_{mej} es la observación del j –ésimo tutorado del e –ésimo estrato en la modalidad m .

N_{me} = número total de docentes del marco muestral, en el e –ésimo estrato, en la modalidad m .

n_{mej}^* = número de docentes efectivamente entrevistados, en el e –ésimo estrato (nivel y grupo de desempeño), en la modalidad m .

s_{me}^2 = es la varianza muestral del e –ésimo estrato en la modalidad m .

$e = 1, 2, \dots, 194$ si $m =$ Educación Básica, y $e = 1, 2, \dots, 60$ si $m =$ Educación Media Superior. En algunos estratos podría no haberse logrado al menos una entrevista, en dichos casos el estrato no contribuye para la estimación del total.

Las estimaciones del error estándar (EE), coeficiente de variación (CV) y efecto de diseño ($deff$) del total \hat{t} con las siguientes expresiones

$$EE(\hat{t}) = \sqrt{\hat{Var}(\hat{t})}$$

$$CV(\hat{t}) = \frac{\sqrt{\hat{Var}(\hat{t})}}{\hat{t}}$$

$$deff = \left(\frac{\hat{Var}(\hat{t})}{\hat{Var}(\hat{t})_{mas}} \right)$$

donde

$\hat{Var}(\hat{t})_{mas}$ =es el estimador de varianza del total \hat{t} bajo muestreo aleatorio simple.

El intervalo de confianza del $100(1 - \alpha)\%$ para el total se construye como

$$\hat{t} \pm z_{\alpha/2} \sqrt{\hat{var}(\hat{t})}$$

$z_{\alpha/2}$ = es el valor del cuantil de una distribución normal estándar que deja un área de $\alpha/2$ a la derecha del él, en particular $\alpha = 0.05$.

Estos cálculos no se realizaron para los estimadores de los parámetros de estudio de los tutores.

Anexo A. PROCESO DE CONSTRUCCIÓN Y DEPURACIÓN DE LOS MARCOS MUESTRALES

A.1 Construcción de los marcos muestrales

Para la construcción de los marcos muestrales de la Encuesta Nacional de Tutoría en Educación Básica y Educación Media Superior se utilizaron las bases de datos de docentes (tutorados) que ingresaron al Servicio Profesional Docente en Educación Básica y Media Superior en el ciclo escolar 2014-2015 y personal (tutores) de Educación Básica y Educación Media Superior que desempeñó la función tutora durante los ciclos 2014-2015 y 2015-2016. La tarea de identificar y validar las variables necesarias para la construcción de los marcos muestrales, para las cuatro poblaciones objetivo, presentó varios obstáculos:

- La ausencia de información de datos de localización de centros de trabajo (plantelos) en Educación Media Superior (municipio, localidad y dirección), tanto para los docentes de nuevo ingreso, como para sus respectivos tutores.
- En Educación Básica, la información de datos de localización de centros de trabajo (plantelos), solamente se presentó para los docentes de nuevo ingreso. En el caso de los tutores, se contó únicamente con las claves de centros de trabajo.
- Se encontraron registros incompletos, incorrectos e inexactos tanto en las bases de datos tutorados como de tutores.

Después del proceso de validación y depuración de las bases de datos, se obtuvieron cuatro marcos muestrales. El resumen de los registros totales válidos se muestran a continuación:

Docente o personal	Modalidad	Nivel Geográfico	Registros
Tutorados	Básica	30 entidades (excepto Oaxaca y Chiapas)	20,647
	Media Superior	30 entidades (excepto Oaxaca y Chiapas)	3,188
Tutores	Básica	27 entidades (excepto Oaxaca, Chiapas, Baja California Sur, Michoacán y Sonora)	8,532
	Media Superior	29 entidades (excepto Oaxaca, Chiapas y Guanajuato)	285

A.2 Validación de la muestra final

Los tamaños de muestra asignados a las cuatro poblaciones objetivos, definidos en la Sección 5, no se alcanzaron debido a la no respuesta. Se llevó a cabo un proceso de validación para que todos los tutorados y tutores entrevistados se encontraran registrados en las bases de datos, con las cuales se diseñaron los marcos muestrales; en casos donde había información faltante de identificación, se buscaron los datos correspondientes en los cuatro marcos muestrales disponibles.

A.3 Ajuste de los tamaños de los estratos

La no respuesta y el proceso de validación de los datos de identificación, de los tutorados y tutores entrevistados, implicaron que algunos estratos no contaran con entrevistas efectivas, por lo tanto, tampoco representación en la muestra. Este hecho se consideró para el cálculo en las estimaciones finales, en la siguiente tabla se describen los tamaños totales que representan los estimadores de los parámetros de estudio.

Docente personal	o	Modalidad	Nivel Geográfico	Registros
Tutorados		Básica	30 entidades (excepto Oaxaca y Chiapas)	20,447
		Media Superior	30 entidades (excepto Oaxaca y Chiapas)	3,188
Tutores		Básica	27 entidades (excepto Oaxaca, Chiapas, Baja California Sur, Michoacán y Sonora)	8,438
		Media Superior	29 entidades (excepto Oaxaca, Chiapas y Guanajuato)	284