



ERCE 2019: Diseño muestral

Oswaldo Palma
INEE-México
Septiembre de 2018



- Dieciocho países participantes
- Tres países en calendario norte y 15 en calendario sur
- Muestras seleccionadas entre el 19 de febrero y el 24 de febrero de 2018 para el calendario norte
- Periodo de selección para el calendario sur del 4 de mayo al 26 de julio de 2018
- Entre 49 y 80 escuelas, dependiendo del promedio de estudiantes en cada grado
- Al menos 905 y hasta 1,351 estudiantes en muestra

- Establecer los métodos que serán utilizados para la selección de unidades de observación para su participación en el estudio y determinar la forma que tendrán los principales estimadores.
- La selección de los métodos busca que los estimadores sean tan precisos como sea posible mientras el costo de la implementación del estudio se mantiene dentro de límites razonables.

- El periodo escolar de referencia para el ERCE-2019 será de **septiembre de 2018 a julio de 2019** para los países que siguen el calendario norte y el comprendido entre **marzo y diciembre de 2019** para los países del calendario sur.

- La cobertura geográfica del estudio estará determinada por los límites territoriales de cada uno de los países participantes salvo las excepciones que en cada caso se señalen en la construcción del marco muestral o durante el desarrollo del operativo.

“... los estudiantes de tercer y sexto grados, dentro del sistema educativo formal, que se consideran aptos para responder pruebas de logro de aprendizaje, y que se encuentran estudiando en escuelas reconocidas por el estado, que contienen aulas que imparten educación correspondiente a los grados objetivo del estudio, en los países que forman parte del LLECE, en total sistemas educativos que aplicaron el estudio.”

UNESCO-OREALC (2016). Reporte Técnico. Tercer Estudio Regional Comparativo y Explicativo, TERCE. Santiago, Chile

- Escuelas donde se imparte *educación no formal*.
- Escuelas para alumnos con necesidades especiales.
- Alumnos en escuelas regulares pero que tienen necesidades especiales.
- Escuelas en donde se imparte educación para adultos.
- Escuelas monolingües en donde se enseña solamente un idioma que no es el español o el portugués en el caso de Brasil.
- Escuelas que se encuentren entre el 2% de escuelas pequeñas en los países del TERCE.

- La unidad de observación y análisis principal para el ERCE-2019 son los estudiantes de grado 03 y 06.
- Otras unidades de observación son:
 - Hogares de los alumnos de grado 03 y 06
 - Padres o tutores de los alumnos
 - Secciones o aulas
 - Maestros que enseñan lenguaje y matemática
 - Directores de las escuelas
 - Escuelas a las que asisten

- Conformado por el listado de escuelas en las que se imparte la educación para los grados que definen la población de interés.
- Deberá incluir todas las escuelas que oficialmente pueden impartir la educación del grado 03 o 06, aun cuando no tengan matrícula registrada para el periodo de referencia.
- También debe incluir las escuelas que pueden satisfacer los criterios de exclusión.

- Cada uno de los países participantes representa un dominio de estudio
- La región conformada por los países participantes es dominio de estudio del ERCE

- En atención a las recomendaciones del CTAN, de la sesión del 9 de febrero de 2018, el diseño muestral del estudio será:
 - Probabilístico
 - Estratificado
 - Por conglomerados
 - En dos etapas (escuela y aula)
 - Con probabilidades diferentes

- Las unidades de muestreo para la primera etapa son las escuelas incluidas en el marco de muestreo
- Las de segunda etapa son las aulas de cada uno de los grados de interés en la escuela durante el periodo de referencia que se obtiene mediante el proceso de validación de datos que cada coordinación nacional realiza.
- El tamaño de cada unidad primaria de muestreo está dado por la suma de la matrícula de grado 03 y la correspondiente a grado 06 incluida en el marco de muestreo.

- El cálculo inicial del tamaño de muestra, n , está dado por

$$n = deff * \left[\frac{1}{\frac{\delta^2}{Z_{\alpha/2}^2 * S^2} + \frac{1}{N}} \right]$$

- Donde
 - S^2 es la varianza de la escala
 - $Z_{\alpha/2}$ es el cuantil de la distribución normal que corresponde al nivel de confianza $(1 - \alpha) * 100\%$
 - δ es el margen de error máximo aceptable para la estimación
 - N es el número de estudiantes del grado de interés en el país
 - $deff$ es el efecto de diseño que refleja el cambio en la eficiencia del estimador cuando se utiliza un diseño diferente al aleatorio simple.

- El número de estudiantes en muestra por país oscila entre 4,000 y 5,500 por grado.
- Para alcanzar el tamaño de muestra del país es necesario seleccionar entre 150 y 300 escuelas
- El número de unidades en muestra es diferente en cada país ya que el valor de los parámetros para su determinación cambia de país a país.

- **Área geográfica**, la cual clasifica cada escuela del marco según el tipo de área donde se ubica como urbana o rural.
 - Los usuarios de las bases de datos deben tener presente que, en general, no es posible la comparación de resultados entre países si estos se presentan por tipo de área geográfica.
- **Dependencia administrativa de la escuela**, que separa a las escuelas en públicas (cuando su administración y financiamiento es público) y privadas cuando su administración es privada.
- **Grados de interés ofertados en la escuela**, que categoriza a las escuelas en grupos según si la escuela ofrece: ambos grados, sólo grado 03 o sólo grado 06.

- El tamaño de muestra será distribuido de manera proporcional al tamaño del estrato.
- La unidad de medida para el tamaño del estrato será la suma de la matrícula en grado 03 y la correspondiente al grado 06 que aparece en el marco de muestreo para la escuela.

- Para el estrato h se seleccionarán n_h escuelas con probabilidad proporcional al número de estudiantes en grado 03 y 06 en la escuela, según la información contenida en el marco.
- Para la selección de las escuelas se utilizará muestreo sistemático.
- En cada una de las n_h escuelas seleccionadas se seleccionará, de acuerdo a la disponibilidad para el periodo de referencia, un aula de cada uno de los grados de interés del estudio.

- Salvo los casos de excepción, todos los estudiantes del aula seleccionada serán seleccionados para la aplicación de los instrumentos.
- La selección de las unidades de cada etapa se realiza de manera independiente.
- El muestreo sistemático no permite construir estimadores insesgados para la varianza de estimadores insesgados de totales o funciones de estos.

- Escuela: Inverso de su probabilidad de selección
- Secciones o aulas: Número de aulas que imparten el grado por el factor de la escuela
- Estudiantes del grado de interés: Factor del aula
- Hogar del alumnos: Factor del estudiante (aula)
- Padres o tutores de los alumnos: Factor del estudiante
- Maestros que enseñan lenguaje, matemática y ciencias: Factor de la escuela o el aula según cuantos se seleccionen
- Directores de las escuelas: Factor de la escuela
- Para cada unidad los factores puede requerir una calibración

- Los estimadores de totales serán los de Horvitz-Thompson
- Por tratarse de un diseño en dos etapas los estimadores de las varianzas de los totales, medias y proporciones; tienen una expresión analítica cerrada por lo que se podrá utilizar dicha expresión para calcularlos.
- Para la estimación de varianzas de otras funciones de totales (p.e. razones o coeficientes de regresión) se puede utilizar el método de linealización o alguno de remuestreo (BRR, Jackknife)

- La disponibilidad de marcos de muestreo define los tiempos de la planeación de cada país
- En general, los del piloto tuvieron buena calidad y han permitido la selección de unidades
- Se sugiere reforzar la atención sobre la construcción y revisión de la variable de población excluida
- Se han identificado otros detalles menores:
 - Códigos intercambiados
 - (1 Urbano, 2 Rural)
 - Información en variables diferentes
 - (sec_2, sec_3, cod1_ubi_esc, cod1_ubi_cen)
 - Tipo de dato diferente para la misma variable
 - (N/S)

GRACIAS POR SU ATENCIÓN

