

2021

**CURSO  
INTERNACIONAL**

**A DISTANCIA**

22 de febrero  
al 30 de abril  
2021



# **DESAGREGACIÓN DE ESTIMACIONES EN ÁREAS PEQUEÑAS USANDO**

## **Información general**

El curso es organizado por la Unidad de Estadísticas Sociales, División de Estadísticas de la Comisión Económica para América Latina y el Caribe (CEPAL) y por el Fondo de Población de las Naciones Unidas (UNFPA).

## ANTECEDENTES

El énfasis de la Agenda 2030 para el Desarrollo Sostenible en no dejar a nadie atrás (“no one left behind”) y alcanzar a los más vulnerables primero presenta a la comunidad estadística tanto desafíos como oportunidades. La desagregación de datos es esencial para comprender si los frutos del desarrollo benefician a todo el espectro de la sociedad, incluidos los más vulnerables. Producir datos desglosados por diversas características de la población y por diferentes niveles de geografía es un desafío para los sistemas estadísticos nacionales, incluidos los de los países desarrollados.. Nuevas y mejores metodologías estadísticas, técnicas innovadoras y un mayor uso de datos administrativos se encuentran entre las iniciativas necesarias para responder a este desafío.

Las encuestas de hogares permiten monitorear continuamente los indicadores a lo largo del tiempo para que los gobiernos puedan adoptar políticas públicas que se centren primero en los más vulnerables. Sin embargo, la precisión de las estimaciones de las encuestas de hogares disminuye cuando se trata de hacer inferencias en subgrupos para los cuales no se diseñó la encuesta. Sin embargo, es posible utilizar las encuestas de hogares como información para producir estimaciones basadas en la información auxiliar disponible que proviene de los censos o registros administrativos.

La División de Estadísticas de la CEPAL y el UNFPA han aunado sus esfuerzos para abordar el problema de la desagregación de estimaciones y para brindar orientación a los países sobre este tema. Como resultado de esta colaboración, el presente curso busca propiciar el desarrollo de capacidades en nuevas técnicas de estimación, que permitan generar estadísticas para aquellos grupos que no están bien representados en una encuesta, a través de la integración de datos provenientes de censos y registros administrativos, mediante la utilización del software estadístico R.

Este curso cuenta con el apoyo financiero del UNFPA y del Programa de Estadísticas y Datos del Décimo Tramo de la Cuenta de Desarrollo de las Naciones Unidas.

# METODOLOGÍA Y ASPECTOS POR CONSIDERAR

## OBJETIVOS GENERALES

El curso permitirá discutir entre los técnicos de los sistemas estadísticos nacionales la utilidad de las encuestas de hogares junto con otras fuentes de datos para producir estadísticas sociales para áreas pequeñas (subgrupos de población específicos). Igualmente, se revisarán las metodologías actuales para la combinación de fuentes de datos para lograr niveles más altos de precisión y compartir los desafíos y las mejores prácticas para cumplir con el llamado de no dejar a nadie atrás.

## OBJETIVOS ESPECÍFICOS

1. Introducir a los asistentes en el paradigma del error total de muestreo como base de análisis de las encuestas de hogares.
2. Repasar las técnicas usuales de ponderación y estimación directa en las encuestas de hogares basadas en el principio de representatividad.
3. Presentar las metodologías de estimación indirecta en el cálculo de estimadores de parámetros desagregados
4. Introducir a los participantes en el paradigma de la inferencia basada en la modelación sobre la población finita.
5. Presentar las metodologías de estimación de Fay-Harriot (modelos de área) y Battese-Harter-Fuller (modelos de unidad)
6. Presentar las metodologías del Banco Mundial ELL y del predictor Empirical Best.
7. Presentar las metodologías asociadas a parámetros complejos que se desvían del supuesto de la normalidad.
8. Realizar prácticas computacionales en R para todas las metodologías de estimación en áreas pequeñas.
9. Realizar mapas de indicadores desagregados a nivel geográfico.

## DIRIGIDO A

Este curso está dirigido al personal técnico de los INE y demás instituciones de los sistemas estadísticos nacionales, así como también a los miembros de organizaciones que realizan investigación social, y a miembros de la academia, como profesores o estudiantes de programas afines al tópico del curso.

## REQUISITOS

Este curso introducirá paulatinamente los conceptos fundamentales de la teoría del muestreo y el ajuste estadístico de modelos mixtos. Se espera que el asistente tenga destreza en el manejo básico de algunos conceptos estadísticos, como medidas de tendencia central, medidas de dispersión y conceptos de probabilidad. De igual forma, el estudiante debe tener conocimientos básicos en el manejo de información sistematizada en hojas de cálculo. El conocimiento y previo manejo del software R es deseable.

## ¿Por qué debo tomar este curso?

- Para entender las diferentes características de la estimación desagregada de los indicadores de los Objetivos de Desarrollo Sostenible.
- Para obtener capacitación en las técnicas modernas de estimación indirecta basada en de hogares, censos y registros administrativos.
- Para aprender a seleccionar la mejor metodología a la hora de obtener estimaciones desagregadas y sus correspondientes errores estándar.
- Para poder implementar un sistema de estimación en áreas pequeñas dentro de los Institutos Nacionales de Estadística en los países.
- Para tener la oportunidad de interactuar directamente con instructores expertos en la estimación de indicadores desagregados y discutir con profesionales que cuentan con bastante experiencia en este tópico.



## Materiales

Debido a la diversidad de herramientas computacionales disponibles para la modelación estadística y estimación en encuestas de hogares, se propone trabajar con el software estadístico **R**, puesto que su acceso es universal, es un software libre y además es multiplataforma (Windows, Mac, Linux). El ILPES de CEPAL proveerá la plataforma virtual de este curso en donde se podrá acceder a los materiales como presentaciones, videos, foros, bases de datos, códigos computacionales, entre otros.



## Duración del curso

El Curso tiene un total de 60 horas académicas, las cuales se desarrollarán del **22 de febrero de 2021 al 30 de abril de 2021**, repartidos en diez módulos. El curso tiene una duración de diez semanas y exige una hora y media diaria de dedicación para lectura de material, visualización de videos y resolución de actividades propuestas.



## Metodología del curso

Las clases serán virtuales a cargo de funcionarios de la División de Estadísticas e instructores internacionales para que los estudiantes puedan replicar los conceptos adquiridos en el curso a través de 10 módulos que repasan los distintos contenidos. Consta de material audiovisual para ejemplificar los aspectos más importantes de las metodologías de estimación en áreas pequeñas y de prácticas computacionales con bases de datos de encuestas de hogares y censos. Así mismo, el progreso del estudiante será valorado por medio de tareas, evaluaciones y un proyecto final.



## CONTENIDOS TEMÁTICOS

0. Objetivos de Desarrollo Sostenible y SAE .
1. Estimadores Directos en encuestas de hogares.
2. Estimadores sintéticos y compuestos.
3. Modelos de área.
4. Modelos de unidad.
5. Mejor predictor empírico.
6. Modelos bayesianos y generalizados.
7. Marco de referencia para la aplicación de modelos SAE.
8. Mapeo de indicadores.
9. Repaso de las metodologías SAE.



## Metodología del curso

El estudiante deberá presentar tres tareas a lo largo del curso, responder cuatro evaluaciones y realizar un proyecto final. Tanto las tareas como el proyecto se deberán subir a la plataforma virtual en las fechas acordadas al inicio del curso y tendrán una calificación que variará de cero a cien puntos. La ponderación de la evaluación es como sigue a continuación:

- Tareas (10% cada una)
- Evaluaciones (10% cada una)
- Proyecto final (30%)

La asistencia a las sesiones del foro virtual es un componente importante en el desarrollo del curso. Es por lo anterior que se requiere que el estudiante participe en al menos 85% de los foros. La nota mínima aprobatoria del curso será de 70 puntos para los estudiantes que acrediten la participación mínima en las sesiones.

## POSTULACIONES

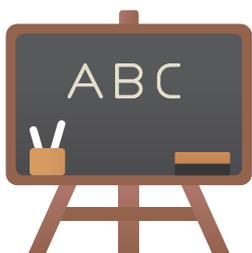
Todos los interesados en postular al Curso deberán completar un formulario de inscripción en línea disponible en el siguiente enlace: <http://sigcaportal.cepal.org>. Una guía práctica sobre cómo postular en este sistema la encontrará haciendo [clic aquí](#).

El interesado que postula por primera vez a un Curso de capacitación de ILPES/CEPAL, primero deberá registrarse en el sistema, adjuntando a su postulación su Currículum Vitae y una carta aval firmada de su institución, reflejando la importancia que para esa Institución tiene la acción formativa del Curso; luego, con su usuario y clave de acceso podrá postular al Curso. En el caso de que el interesado ya está registrado, sólo debe ingresar con su usuario y clave de acceso, luego postular al curso. Eso marca el inicio del proceso de selección de las candidaturas que se recibirán hasta el **jueves 31 de enero de 2021**, las aplicaciones recibidas después de esta fecha no serán consideradas en el proceso de selección. El comité enviará a los candidatos seleccionados un correo electrónico de confirmación. Quienes sean seleccionados, deberán confirmar su participación en el curso.

# TEMARIO POR MODULOS

Módulo	Título del tema
<b>Módulo 0 - Objetivos de Desarrollo Sostenible y SAE</b> Andrés Gutiérrez 22 al 26 de febrero de 2021	Objetivos, metas e indicadores No dejar a nadie atrás Coeficiente de variación Criterios de calidad Áreas pequeñas
<b>Módulo 1 - Estimadores Directos</b> Pedro Silva 1 al 5 de marzo de 2021	Introducción Indicadores comunes de pobreza Métodos directos para la desagregación de datos de pobreza Estimador tipo Hájek Estimadores GREG y de calibración Aplicación
<b>Módulo 2 - Estimadores sintéticos</b> Diego Lemus 8 al 12 de marzo de 2021	Introducción Estimador postestratificado sintético Estimador de regresión a nivel de área (I) Estimador de regresión a nivel de área (II) Estimador de regresión a nivel de individuo Estimadores compuestos Aplicación
<b>Módulo 3 - Modelos de área</b> Cristian Tellez 15 al 19 de marzo de 2021	Introducción Efectos aleatorios Error cuadrado medio Ventajas y desventajas Aplicación
<b>Módulo 4 - Modelos de unidad</b> Diego Lemus 22 al 26 de marzo de 2021	Introducción Modelos con errores anidados Modelo de Battese-Harter-Fuller (I) Modelo de Battese-Harter-Fuller (II) Modelo de Elbers-Lanjouw-Lanjouw Aplicación
<b>Módulo 5 - Mejor predictor empírico</b> Ángela Luna 29 al 31 de marzo de 2021	Introducción Modelo con errores anidados Cálculo del predictor empírico Estimación del error cuadrático medio Aplicación
<b>Módulo 6 - Modelos bayesianos y generalizados</b> Andrés Gutiérrez 5 al 9 de abril de 2021	Introducción Métodos bayesianos (I) Métodos bayesianos (II) Modelos lineales generalizados (I) Modelos lineales generalizados (II) Aplicación
<b>Módulo 7 - Marco de referencia</b> Natalia Rojas 12 al 16 de abril de 2021	Especificación Análisis y adaptación Transformaciones Evaluación Resumen Emdí
<b>Módulo 8 - Mapas</b> Felipe Molina 19 al 23 de abril de 2021	Datos cartográficos Introducción a ggplot2 Mapas en ggplot2 Mapas en tmap Mapas en leaflet
<b>Módulo 9 - Repaso de las metodologías SAE</b> Isabel Molina 26 al 30 de abril de 2021	Introducción Estimadores directos Estimadores indirectos Modelos de área Modelos de unidad Modelos lineales generalizados Extensiones

# INSTRUCTORES



El personal docente del curso está integrado por personal de CEPAL, así como por profesores con amplio reconocimiento internacional y consultores altamente calificados. A continuación, se presentan los instructores del curso

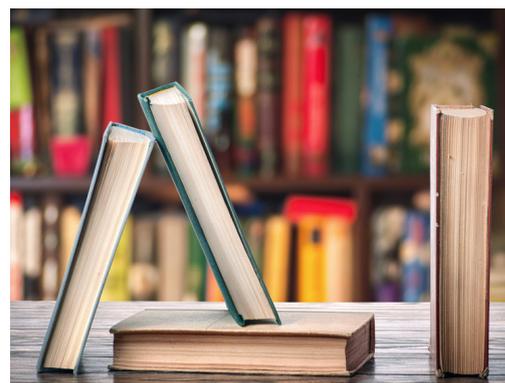
- PhD. Isabel Molina. Universidad Carlos Tercero de Madrid (España).
- PhD. Pedro Silva. ENCE/IBGE (Brasil).
- PhD. Ángela Luna. Universidad de Southampton (Inglaterra).
- PhD. Natalia Rojas. Universidad Libre de Berlín (Alemania).
- PhD. Cristian Tellez. Universidad Santo Tomás (Colombia).
- MSc. Diego Lemus. DNP (Colombia).
- PhD. Andrés Gutiérrez. UN-CEPAL.
- MSc. Felipe Molina. UN-CEPAL.



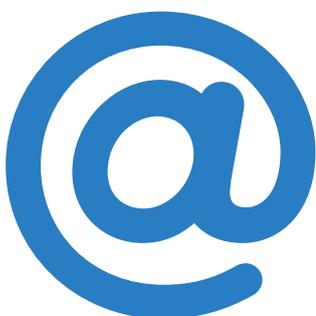
# BIBLIOGRAFÍA

A continuación, se indican los textos que se utilizarán en el desarrollo de este curso:

- Molina, I. (2018). *Desagregación de datos en encuesta de hogares: metodologías de estimación en áreas pequeñas*. CEPAL.
- Rao, J, & Molina, I. (2015) *Small Area Estimation*. Wiley.
- Gutiérrez, H. A. (2016). *Estrategias de muestreo: diseño de encuestas y estimación de parámetros* (Segunda edición). Ediciones de la U.



# CONTACTO



## COORDINACIÓN ACADÉMICA

Para información sobre el contenido del curso contactar a:

**Sr. Andrés Gutiérrez**

Email: [andres.gutierrez@cepal.org](mailto:andres.gutierrez@cepal.org)

Teléfono: 56-2-2210-2665

## COORDINACIÓN ADMINISTRATIVA

Para información general sobre el curso o aspectos logísticos, contactar a:

**Sr. Juan Alvez**

Email: [juan.alvez@cepal.org](mailto:juan.alvez@cepal.org)

Teléfono: 56-2-2210-2474