

Estimación en áreas pequeñas

Introducción a la estimación en dominios de interés

Andrés Gutiérrez, Ph.D.

CEPAL - Unidad de Estadísticas Sociales

- 1 *¿Por qué utilizar los métodos SAE?*
- 2 *¿Qué es un área pequeña?*
- 3 *Consideraciones adicionales*
- 4 *Algunas aplicaciones actuales*

¿Por qué utilizar los métodos SAE?

Justificación

- Los estimadores directos, basados solo en unidades de muestreo observadas para cada área pequeña, no son suficientemente confiables.

Justificación

- Los estimadores directos, basados solo en unidades de muestreo observadas para cada área pequeña, no son suficientemente confiables.
- Tamaño de muestra pequeño o incluso ninguna unidad observada (falta de información).

Justificación

- Los estimadores directos, basados solo en unidades de muestreo observadas para cada área pequeña, no son suficientemente confiables.
- Tamaño de muestra pequeño o incluso ninguna unidad observada (falta de información).
- El coeficiente de variación (CV) es demasiado alto para el indicador objetivo a nivel de área.

Justificación

Coeficiente de variación (%)	Número de Observaciones	
	Bajo	Alto
[20 , 100]	Estimador no confiable	Estimador no confiable
[15 , 20)	Estimador no confiable	Descriptivo
[5 , 15)	Descriptivo	Estimador confiable
(0 , 5)	Estimador confiable	Estimador confiable

Figura1: Confiabilidad de las estimaciones

Justificación

Área	Obs	Mean	Std. Dev.	Min	Max
region	15	12,0	4,9	5,5	22,5
provincia	26	19,5	7,1	9,8	35,5
ciudad	33	20,0	6,5	6,0	34,9
comuna	237	34,8	26,0	0,0	108,7

Figura2: Coeficientes de variación en Chile

Justificación

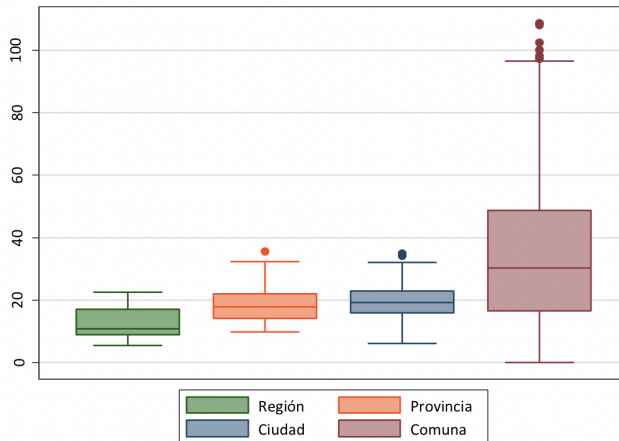


Figura3: Distribución de los coeficientes de variación en Chile

Justificación

Cuando los estimadores directos no son confiables para algunos dominios de interés, existen dos opciones:

- 1 Sobremuestreo: aumentar el tamaño de la muestra en los dominios de interés (aumento de los costos).

Justificación

Cuando los estimadores directos no son confiables para algunos dominios de interés, existen dos opciones:

- 1 Sobremuestreo: aumentar el tamaño de la muestra en los dominios de interés (aumento de los costos).
- 2 Aplicar técnicas estadísticas que permitan estimaciones confiables en esos dominios, métodos SAE.

Justificación

- Durante la última década ha habido una demanda creciente de (objetivo y subjetivo) indicadores de progreso y bienestar.

Justificación

- Durante la última década ha habido una demanda creciente de (objetivo y subjetivo) indicadores de progreso y bienestar.
- Estas medidas desempeñan un papel central para los responsables de las políticas, para planificar y verificar la efectividad de las mismas.

Ejemplos

- Indicadores de pobreza: en riesgo de pobreza, ingreso de los hogares.

Ejemplos

- Indicadores de pobreza: en riesgo de pobreza, ingreso de los hogares.
- Indicadores del mercado de trabajo: Tasa de desempleo, Satisfacción con el trabajo, etc.

Ejemplos

- Indicadores de pobreza: en riesgo de pobreza, ingreso de los hogares.
- Indicadores del mercado de trabajo: Tasa de desempleo, Satisfacción con el trabajo, etc.
- Indicadores de salud: esperanza de vida media, porcentaje de población con conductas peligrosas (obesidad, fumadores, etc.)

Justificación

- Para ser informativos y efectivos, estos indicadores deben elegirse a el nivel apropiado de desagregación.

Justificación

- Para ser informativos y efectivos, estos indicadores deben elegirse a el nivel apropiado de desagregación.
- Los indicadores pueden ser desagregados a lo largo de varias dimensiones, incluyendo áreas geográficas, grupos demográficos, grupos de ingresos / consumo y grupos sociales.

¿Qué es un área pequeña?

¿Qué es un área pequeña?

- La mayoría de las encuestas nacionales están planificadas para entregar estimaciones confiables a nivel nacional y regional pero a niveles más bajos se reduce la precisión.

¿Qué es un área pequeña?

- La mayoría de las encuestas nacionales están planificadas para entregar estimaciones confiables a nivel nacional y regional pero a niveles más bajos se reduce la precisión.
- Un área pequeña es un dominio para el cual el tamaño de muestra específico no es suficientemente grande para obtener estimaciones confiables.

¿Qué es un área pequeña?

- La mayoría de las encuestas nacionales están planificadas para entregar estimaciones confiables a nivel nacional y regional pero a niveles más bajos se reduce la precisión.
- Un área pequeña es un dominio para el cual el tamaño de muestra específico no es suficientemente grande para obtener estimaciones confiables.
- Habitualmente son dominios no planificados y su tamaño de muestra esperado es aleatorio y es más grande a medida que aumenta el tamaño de la población del área.

¿Qué es un área pequeña?

La subpoblación de interés puede ser una zona geográfica o subgrupos socioeconómicos.

- Geográfico: provincias, áreas del mercado de trabajo, municipios, sectores censales para medir por ejemplo la tasa de desempleo a nivel comunal.

¿Qué es un área pequeña?

La subpoblación de interés puede ser una zona geográfica o subgrupos socioeconómicos.

- Geográfico: provincias, áreas del mercado de trabajo, municipios, sectores censales para medir por ejemplo la tasa de desempleo a nivel comunal.
- Dominio de subgrupos específicos: edad \times sexo \times raza dentro del ámbito geográfico de una zona, para medir por ejemplo la tasa de desempleo por sexo o edad específica en las zonas urbanas.

¿Qué es un área pequeña?

- La solución es **tomar prestada fuerza** de otras áreas y/o en diferentes ocasiones mediante modelos explícitos o implícitos que explotan la relación entre variables aumentando el tamaño efectivo de la muestra.

¿Qué es un área pequeña?

- La solución es **tomar prestada fuerza** de otras áreas y/o en diferentes ocasiones mediante modelos explícitos o implícitos que explotan la relación entre variables aumentando el tamaño efectivo de la muestra.
- El modelo proporciona un enlace a áreas relacionadas y/o períodos de tiempo a través de información complementaria tales como recuentos de censos (recientes o actuales) o registros administrativos relacionados con la variable objetivo.

Algunos métodos

- Estimador sintético: En el contexto de subpoblaciones, los estimadores se llaman sintéticos cuando éstos se basan en un estimador directo y se estiman a partir de información auxiliar a través de un modelo.

Algunos métodos

- Estimador sintético: En el contexto de subpoblaciones, los estimadores se llaman sintéticos cuando éstos se basan en un estimador directo y se estiman a partir de información auxiliar a través de un modelo.
- Estimador compuesto: es una combinación lineal entre un estimador directo y un estimador sintético. Representa un buen compromiso entre las características de los dos componentes.

Algunos métodos

- El estimador compuesto está dado por una combinación lineal de estimador sintético y estimador directo equilibrando el sesgo potencial del estimador sintético contra la inestabilidad del estimador directo (compensación entre precisión y sesgo).

Algunos métodos

- El estimador compuesto está dado por una combinación lineal de estimador sintético y estimador directo equilibrando el sesgo potencial del estimador sintético contra la inestabilidad del estimador directo (compensación entre precisión y sesgo).
- Las estimaciones más grandes de áreas pequeñas están más cerca de las estimaciones directas mientras que las más pequeñas están más cerca de las estimaciones sintéticas.

Algunos métodos

- Los estimadores SAE se dividen en dos tipos principales dependiendo de cómo se aplican los modelos a los datos dentro de las áreas pequeñas: nivel de área y nivel de unidad.

Algunos métodos

- Los estimadores SAE se dividen en dos tipos principales dependiendo de cómo se aplican los modelos a los datos dentro de las áreas pequeñas: nivel de área y nivel de unidad.
- Los estimadores de área pequeña se basan en cálculos de nivel de área si los modelos vinculan la variable de interés y con variables auxiliares x específicas del área.

Algunos métodos

- Se llaman modelos a nivel de unidad si se vinculan valores individuales para las variables auxiliares específicas de la unidad.

Algunos métodos

- Se llaman modelos a nivel de unidad si se vinculan valores individuales para las variables auxiliares específicas de la unidad.
- Los estimadores basados en áreas pequeñas se calculan a nivel de área si los datos de la unidad no están disponibles.

Algunos métodos

- Se llaman modelos a nivel de unidad si se vinculan valores individuales para las variables auxiliares específicas de la unidad.
- Los estimadores basados en áreas pequeñas se calculan a nivel de área si los datos de la unidad no están disponibles.
- También pueden ser calculados si los datos de nivel de unidad están disponibles resumiéndolos en el nivel de área apropiado.

Consideraciones adicionales

Proceso de estimación

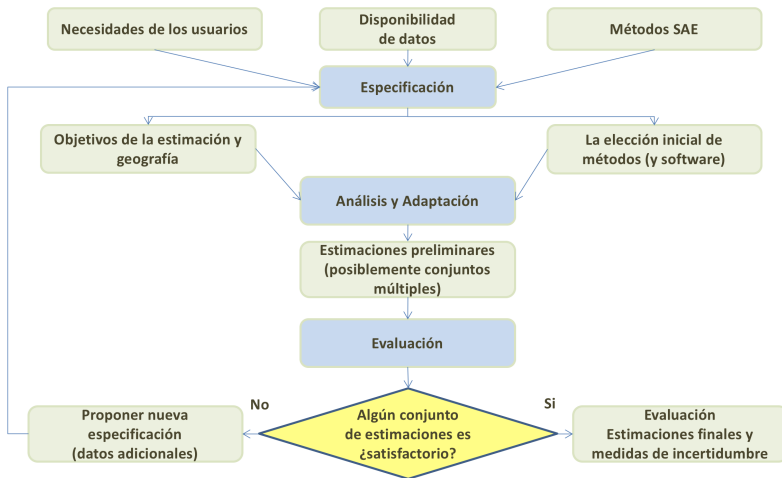


Figura4: Producción de estadísticas con SAE

Algunos riesgos

La producción de estimaciones en área pequeña involucra riesgos que se deben tener en consideración:

- El tamaño de las áreas pequeñas en los términos del número de unidades que les pertenecen es también una consideración importante. Áreas que son demasiado pequeñas pueden presentar problemas de confidencialidad.

Algunos riesgos

La producción de estimaciones en área pequeña involucra riesgos que se deben tener en consideración:

- El tamaño de las áreas pequeñas en los términos del número de unidades que les pertenecen es también una consideración importante. Áreas que son demasiado pequeñas pueden presentar problemas de confidencialidad.
- Las estimaciones de área pequeña pueden diferir demasiado de las estadísticas basadas en el conocimiento local.

Algunos riesgos

La producción de estimaciones en área pequeña involucra riesgos que se deben tener en consideración:

- El tamaño de las áreas pequeñas en los términos del número de unidades que les pertenecen es también una consideración importante. Áreas que son demasiado pequeñas pueden presentar problemas de confidencialidad.
- Las estimaciones de área pequeña pueden diferir demasiado de las estadísticas basadas en el conocimiento local.
- Las fuentes de información y el diseño utilizado pueden ser no sustentables en el tiempo.

Algunos riesgos

- El compromiso y la voluntad de la agencia para apoyar estas metodologías a través de sistemas y personal capacitado en la materia.

Algunos riesgos

- El compromiso y la voluntad de la agencia para apoyar estas metodologías a través de sistemas y personal capacitado en la materia.
- Disponibilidad de datos auxiliares correlacionados con la variable de interés.

Algunos riesgos

- El compromiso y la voluntad de la agencia para apoyar estas metodologías a través de sistemas y personal capacitado en la materia.
- Disponibilidad de datos auxiliares correlacionados con la variable de interés.
- Tamaño de muestra de debe ser suficientemente grande para permitir estimaciones confiables mediante el uso de los datos de la encuesta y los datos auxiliares existentes.

Consideraciones

- Todos los métodos SAE requieren datos auxiliares a nivel del área pequeña desde el cual **toman prestada la fuerza**.

Consideraciones

- Todos los métodos SAE requieren datos auxiliares a nivel del área pequeña desde el cual **toman prestada la fuerza**.
- La efectividad de los métodos SAE depende del grado de asociación entre la variable de interés y los datos auxiliares.

Consideraciones

- Todos los métodos SAE requieren datos auxiliares a nivel del área pequeña desde el cual **toman prestada la fuerza**.
- La efectividad de los métodos SAE depende del grado de asociación entre la variable de interés y los datos auxiliares.
- La búsqueda de buenas variables auxiliares es crítica, incluida la construcción imaginativa de tales variables.

Consideraciones

- Todos los métodos SAE requieren datos auxiliares a nivel del área pequeña desde el cual **toman prestada la fuerza**.
- La efectividad de los métodos SAE depende del grado de asociación entre la variable de interés y los datos auxiliares.
- La búsqueda de buenas variables auxiliares es crítica, incluida la construcción imaginativa de tales variables.
- Los datos auxiliares deben medirse de manera consistente a través de las áreas pequeñas, pero pueden incluir estimaciones de muestras grandes con error de muestreo conocido.

Desafíos

- Aumento de las tasas de no respuesta.

Desafíos

- Aumento de las tasas de no respuesta.
- Aumento de costos, menos financiación.

Desafíos

- Aumento de las tasas de no respuesta.
- Aumento de costos, menos financiación.
- Aumento de la demanda de estimaciones para dominios pequeños como por raza, etnia o pobreza.

Desafíos

- Aumento de las tasas de no respuesta.
- Aumento de costos, menos financiación.
- Aumento de la demanda de estimaciones para dominios pequeños como por raza, etnia o pobreza.
- Aumento de la demanda de estimaciones de áreas pequeñas.

Desafíos

- Aumento de las tasas de no respuesta.
- Aumento de costos, menos financiación.
- Aumento de la demanda de estimaciones para dominios pequeños como por raza, etnia o pobreza.
- Aumento de la demanda de estimaciones de áreas pequeñas.
- Aumento de la complejidad en los contenidos de los cuestionarios y por lo tanto la carga de respuesta.

Desafíos

- Aumento de las tasas de no respuesta.
- Aumento de costos, menos financiación.
- Aumento de la demanda de estimaciones para dominios pequeños como por raza, etnia o pobreza.
- Aumento de la demanda de estimaciones de áreas pequeñas.
- Aumento de la complejidad en los contenidos de los cuestionarios y por lo tanto la carga de respuesta.
- Aumento de la demanda de análisis secundarios, uso público y archivos de datos de uso restringido.

Algunas aplicaciones actuales

Mapas de pobreza

Estimados por el Banco Mundial en diversos países: Perú, Brasil, Guatemala, Nicaragua, Panamá, Ecuador. Combinan datos provenientes de encuestas con datos censales.

Small Area Income and Poverty Estimates (SAIPE)

El Small Area Income and Poverty Estimates (SAIPE) es un programa del Census Bureau de los Estados Unidos. El programa SAIPE produce estimaciones de pobreza para el total de la población y la media de los ingresos por hogares estimada anualmente para todos los condados y estados.

Estimaciones mensuales de empleo local y estatal

A través del programa de Estadísticas de desempleo del área local, la BLS produce estimaciones mensuales del empleo y desempleo total para aproximadamente 7300 áreas, incluido regiones censales, divisiones, estados, condados y ciudades.

Las estimaciones son basadas en datos de varias fuentes, incluyendo el CPS, el programa de estadísticas de empleo actual (CES), los sistemas de seguro de desempleo de los estados (UI) y el censo decenal.

La experiencia canadiense

Utilizaron una serie de fuentes primarias de datos: Censo Canadiense de población y vivienda, La encuesta de fuerza laboral (LFS) Canadiense y el sistema de seguro gubernamental federal de desempleo (UI). Otros datos de áreas pequeñas de mercado laboral fueron utilizados como información auxiliar.

Dado que el objetivo principal era minimizar el error de estimación del modelo, probaron tres técnicas de estimación: Estimación sintética, SPREE y regresión.

En el Reino Unido

En conjunto con la universidad de Southampton, se realizó una estimación del desempleo basado en la definición de la OIT. Se exploraron distintos enfoques utilizando la alta correlación entre las estimaciones de desempleo de la LFS y el número de demandantes de de beneficios de subsidios para solicitantes de empleo.

El enfoque principal fue desarrollar modelos de regresión que vinculen las estimaciones de desempleo con información de los solicitantes otorgada por las autoridades locales de los distritos (ONS, 2001b).

En Italia

La experiencia de Italia fue un intento de mejorar el rendimiento de la estimación de la tasa de desocupación a nivel sub-regional a través de la LFS del ITSTAT (Italy statistics) utilizando como referencia los dominios que abarcan los estratos que se pueden agregar a los municipios.

La tasa de desempleo se estimó utilizando un modelo lineal mixto con efectos de área espacialmente correlacionados y covariables como el sexo, edad y el desempleo a nivel área del censo anterior.

¡Gracias!

Email: andres.gutierrez@cepal.org