

Indicadores sociodemográficos a nivel área pequeña para la Encuesta Intercensal 2015: Incorporación de efectos espaciales y temporales

Rogelio Ramos Quiroga

`rramosq@cimat.mx`

4o Seminario Sobre Investigaciones Financiadas por el
Fondo Sectorial CONACYT-INEG
Cd. de México, 28 de Febrero de 2017

Plan de la Presentación

- Descripción general de la EIC 2015
- Demanda del Fondo Sectorial CONACY-INEGI 2015-1
- Propuesta
 - Fuentes de información
 - Variables a estudiar
 - Modelos básicos
 - Retos y compromisos
- Equipo de trabajo
- Referencias

Encuesta Intercensal 2015

La Encuesta Intercensal 2015 (EIC)

Los censos decenales de población y vivienda del INEGI son la fuente más completa de información con la que contamos. Dada la demanda de información actualizada en periodos más cortos, el INEGI ha organizado conteos y encuestas intercensales en 1995, 2005 y 2015.

- Encuesta con cobertura temática amplia.
 - Características de las Viviendas
 - Características de la Población.
 - Fecundidad
 - Mortalidad infantil
 - Ingreso por trabajo
 - Conformación de Hogares.
- La EIC 2015 incluyó cerca de 6 millones de viviendas.
 - Esquema de muestreo estratificado por tamaño de localidad y nivel socioeconómico. Se seleccionaron áreas geográficas completas usando muestreo aleatorio simple, visitándose todas las viviendas.

Demanda 1: Fondo Sectorial CONACYT-INEGI 2015-1

Demanda 1. FONSEC CONACYT-INEGI 2015-1

- Nombre del tema
 - Medición de características sociodemográficas mediante las técnicas de estimación de áreas pequeñas, para la Encuesta Intercensal 2015.
- Objetivos específicos
 - Desarrollar indicadores de la dinámica demográfica en los temas de fecundidad, mortalidad infantil, sector de actividad, ocupación y otros ingresos; mediante métodos indirectos con la Encuesta Intercensal 2015, que permitan apoyar la información proporcionada mediante las técnicas de estimación de áreas pequeñas. Obtener una metodología que permita medir las características sociodemográficas ya mencionadas a nivel municipal, para cubrir las demandas orientadas a la construcción de políticas y proyectos de desarrollo social; y que no se puedan estimar directamente con datos de la encuesta intercensal 2015, utilizando las técnicas de estimación de áreas pequeñas.

Proyecto Propuesto

Propuesta de Trabajo

- Proponemos estudiar, producir modelos y presentar estimaciones a nivel municipal basados en las técnicas indirectas de estimación de áreas pequeñas de:
 - Tasas de mortalidad infantil
 - Tasas de fecundidad
 - Ingreso por trabajo
- Restringiéndonos a los estados:
 - Aguascalientes
 - Guerrero
 - Michoacán
 - Tamaulipas

Fuentes de Información

Preparación de Bases de Datos. Los modelos estadísticos que se propondrán serán construidos en base a diferentes fuentes de información. Entre ellas destacamos las más relevantes para estudiar fecundidad, mortalidad infantil e ingreso:

- Encuestas Intercensales 2015, 2005, 1995
- Censos de Población y Vivienda 2010, 2000
- ENIGH (Encuesta Nacional de Ingresos y Gastos de los Hogares)
- SIMBAD (Sistema Municipal de Bases de Datos)
- ENADID (Encuesta Nacional de la Dinámica Demográfica)

De éstas encuestas y sistemas planeamos construir variables auxiliares a nivel municipal, tanto a nivel contemporáneo como con variables con rezagos temporales, además de información georeferenciada que nos permitirá incorporar efectos espaciales.

Variables a Estudiar 1

- Fecundidad
 - Tasa Global de Fecundidad (TGF)
 - Tasas por grupos de edad
- Mortalidad
 - Porcentaje de hijos fallecidos por grupos de edad de las madres
- Ingreso por Trabajo de la Población Ocupada
- **Covariables a considerar**
 - Educación, acceso a servicios de salud, alimentación, acceso al agua, etnicidad, características de la vivienda, etc.

Variables a Estudiar 2

Formalmente, las cantidades objetivo son:

- **Tasa de mortalidad infantil.** Cociente de las defunciones de menores de un año de edad ocurridas en un periodo determinado, entre los hijos nacidos vivos durante el mismo lapso, multiplicado por mil.
- **Tasa específica de fecundidad.** Cociente del número de nacimientos ocurridos en las mujeres de determinado grupo quinquenal de edad, con respecto de la población femenina de ese mismo grupo de edad.
- **Tasa global de fecundidad.** Número promedio de hijos que tendría una mujer al término de su vida fértil si a lo largo de ese periodo estuviera sujeta a las tasas específicas de fecundidad del año considerado.
- **Ingresos por trabajo.** Percepción monetaria que la población ocupada de 12 y más años de edad obtiene o recibe del trabajo que desempeñó en la semana de referencia.

Variables a Estudiar 3

Preguntas relevantes de la EIC 2015

- Ingreso por Trabajo de la Población Ocupada
 - (34) ¿Cuánto gana LA PERSONA por ese trabajo?
- Fecundidad y Mortalidad
 - (42) En total, ¿cuántas hijas e hijos que nacieron vivos ha tenido LA PERSONA?
 - (43) De las hijas e hijos que nacieron vivos de LA PERSONA, ¿cuántos han muerto?
 - (44) ¿Cuántas de las hijas e hijos de LA PERSONA viven actualmente?
 - (45) ¿En qué mes y año nació la última hija o hijo nacido vivo de LA PERSONA?
 - (46) Esta última hija o hijo de LA PERSONA, ¿vive actualmente?
 - (47) Si no, ¿qué edad tenía cuando murió?

Fuentes de Información, EIC

Número de registros disponibles en la EIC 2015

Estado	Población	Vivienda
Aguascalientes	177,853	43,613
Guerrero	1,144,440	279,157
Michoacán	1,133,370	291,433
Tamaulipas	400,777	116,613

Número de variables por temática en la EIC 2015

Tema	No. de Variables
Fecundidad y Mortalidad	6
Características Económicas	15

Modelos Básicos

Los modelos básicos son los modelos lineales o lineales generalizados con efectos mixtos (para e.g. ingreso promedio o tasas, respectivamente). Para ingreso, consideraremos, en primera instancia, modelos de la forma

$$\log(y_{ijk}) = x_{ij}^T \beta + v_i + u_{ij} + e_{ijk}$$

el índice i se refiere a dominio de área pequeña (e.g. municipio), j indica la UPM y k la vivienda. Asumimos que las covariables (variables auxiliares) sólo se pueden asociar con las respuestas a nivel UPM. Los efectos v_i pueden o no ser aleatorios, dependiendo del nivel de desagregación requerido: Fijos, si el nivel es municipal, o aleatorios si el nivel es localidad.

Jerarquía de Modelos

Proponemos considerar una jerarquía de modelos:

- Estimaciones directas
- Estimaciones indirectas basadas en modelos SAE (con covariables contemporáneas o casi contemporáneas)
- Incorporación de efectos espaciales (explorando alternativas de definición de cercanía entre municipios)
- Incorporación de efectos temporales (modelos para series cortas)
- Incorporación de ambos tipos de efectos

Modelación de las Variables de Interés

Una vez estructurada la información relevante decidiremos si las variables de interés se tomarán directamente de la EIC 2015 o si prepararemos índices apropiados como proxis de esas variables. Los modelos propuestos, en principio, consistirán de modelos de nivel área para las tasas de mortalidad infantil y fecundidad, y modelos de nivel unidad para ingreso. Se adaptarán las metodologías propuestas en González-Manteiga et al. (2008) y en Marhuenda et al. (2013) para las condiciones de nuestra problemática, así como también se incorporarán los factores de expansión apropiados en todos los procesos de estimación y predicción correspondientes. Se hará uso extensivo del lenguaje de programación R.

Retos Metodológicos y Compromisos

- Comparación de errores cuadráticos medios usando aproximaciones analíticas versus metodologías basadas en remuestreo (bootstrap)
- Comparación SAE versus estimadores directos
- Selección de modelos
- Incorporación de efectos del diseño muestral en los ajustes de los modelos
- Interacción con expertos en temas demográficos para el tratamiento adecuado de variables demográficas
- Producción de mapas. Herramientas de visualización de resultados
- Documentación y Capacitación

Equipo de Trabajo

Equipo de Trabajo

- Georges Bucyibaruta (EST)
- Francisco Javier Rivera Ramírez (EST)
- Rogelio Ramos Quiroga (EST)
- Edgar Jiménez Peña (COMP)
- José Bertín Ramírez (ECO)
- Andrés Carreón Lozano (ECO)

Referencias

Referencias

- **Battese, G.E., Harter, R.M, and Fuller, W.A.** (1988). An error-components model for prediction of county crop areas using survey and satellite data. *Journal of the American Statistical Association*, Vol. 83, No. 401, 28-36
- **Benavent, R. and Morales, D.** (2016). Multivariate Fay–Herriot models for small area estimation. *Computational Statistics and Data Analysis*. 94, 372–390.
- **Coelho, P.S. and Pereira, L.N.** (2011). A spatial unit level model for small area estimation. *Statistical Journal*. Vol. 9, No. 2, 155–180.
- **Encuesta Intercensal** (2015). Encuesta Intercensal 2015: Síntesis metodológica y conceptual. Instituto Nacional de Estadística y Geografía. INEGI
- **Fay, R.E. and Herriot, R.A.** (1979). Estimates of income for small places: An application of James-Stein procedures to census data. *Journal of the American Statistical Association*, Vol. 74, No. 366, 269-277

Referencias

- **González-Manteiga, W., Lombardía, M.J., Molina, I., D. Morales, D. and Santamaría, L.** (2008). Analytic and bootstrap approximations of prediction errors under a multivariate Fay–Herriot model. *Computational Statistics and Data Analysis*, 52, 5242–5252
- **Maruhenda, Y., Molina, I. and Morales, D.** (2013). Small area estimation with spatio-temporal Fay–Herriot models. *Computational Statistics and Data Analysis*, 58, 308–325
- **Rao, J.N.K.** (2003). *Small Area Estimation*. Wiley.
- **Suárez Campos, M.A.** (2010). Estimación del ingreso promedio por vivienda en los municipios del Estado de Sonora: Una aplicación de la estimación para áreas pequeñas. Tesis de Maestría en Estadística Oficial, Programa CIMAT-INEGI.
- **Suárez Campos, M.A., Aguilar Mata, G. y Mejía González, R.** (2015). Estimación del ingreso por trabajo en los municipios y delegaciones de México utilizando técnicas de estimación para áreas pequeñas. *Revista Internacional de Estadística y Geografía*, Vol. 6, No. 3, págs. 44-61.

GRACIAS!!