

Estimación de indicadores de pobreza con modelos Bayesianos de área pequeña

Luis E. Nieto Barajas

Departamento de Estadística
ITAM

Seminario INEGI – 28 Nov, 2023



Contents

- 1 Introducción
- 2 Modelo municipal
- 3 Modelo individual
- 4 Conclusiones

Pobreza multidimensional

- Pobreza multidimensional (7) : ingreso + 6 rezagos
 - ic-rezedu : Rezago educativo
 - ic-asalud : Acceso a servicios de salud
 - ic-segsoc : Acceso a seguridad social
 - ic-cv : Calidad y espacio de vivienda
 - ic-sbv : Servicios básicos de vivienda
 - ic-ali : Acceso a alimentación nutritiva y de calidad

 - ict : Ingreso corriente total

- CONEVAL : produce indicadores de pobreza multidimensional cada 2 años a nivel estatal y cada 5 años a nivel municipal

Fuentes

- ENIGH : Encuesta Nacional de Ingreso y Gasto en los Hogares
- MCS : Módulo de Condiciones Socioeconómicas
- MCS-ENIGH tiene representatividad nacional y estatal, pero no municipal
- Censo : Censo de población y vivienda
- Censo ampliado : Contiene variables para medir pobreza multidimensional

Indicadores

- Indicadores oficiales de pobreza se obtienen de MCS-ENIGH
 - Indicadores nacionales y estatales
- Indicadores municipales : se obtienen del MCS-ENIGH para los municipios “grandes” y se estiman con modelos de EAP en los municipios “pequeños”
- Del Censo ampliado se pueden obtener 4 de los 7 indicadores y se tiene representatividad municipal, los otros 3 se obtienen de modelos de EAP

Idea

- Definir los 7 indicadores de pobreza $\mathbf{Y}_i = (Y_{1i}, Y_{2i}, \dots, Y_{7i})$ para los municipios grandes usando MCS-ENIGH
- Definir los 7 indicadores de pobreza $\mathbf{X}_i = (X_{1i}, X_{2i}, \dots, X_{7i})$ para todos los municipios usando el censo ampliado
- Calibrar los indicadores obtenidos del Censo ampliado usando modelos de EAP
- Predecir los indicadores de pobreza multidimensional \mathbf{Y}_i para los municipios pequeños

Modelo estadístico

- Modelo multivariado : $\mathbf{Y}_i \rightarrow \mathbf{W}_i$ (transformación logística)

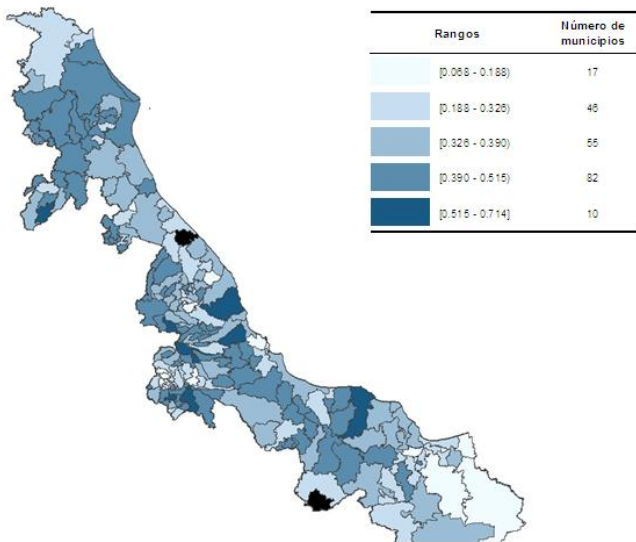
$$\mathbf{W}_i | \theta_i, \Sigma \stackrel{\text{ind}}{\sim} \mathbf{N}_m(\theta_i, c_i \Sigma), \quad i = 1, \dots, n$$

- Σ matriz de var-cov y c_i un factor de escala (diseño)
- Media

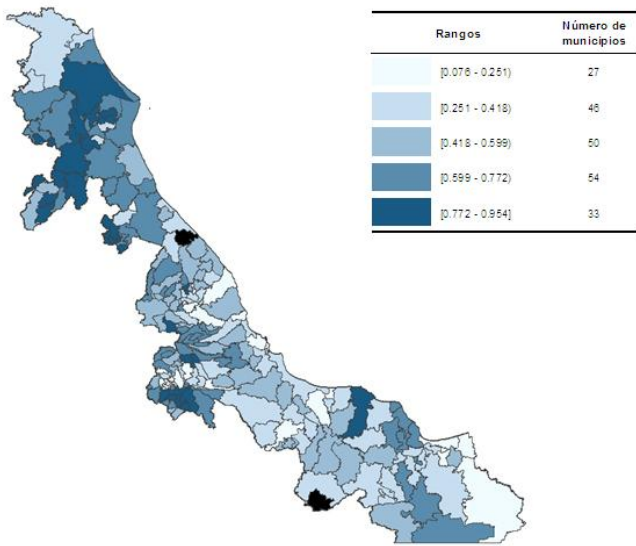
$$\theta_i = B\mathbf{x}_i + v_i \mathbf{1} + u_i \mathbf{1}$$

- $\mathbf{1}$ vector de unos
- Efectos fijos $B\mathbf{x}_i$ con B matriz de coeficientes
- Efectos individuales $v_i \stackrel{\text{iid}}{\sim} \mathbf{N}(0, \sigma_v^2)$
- Efectos espaciales $\mathbf{u}' = (u_1, \dots, u_n) \sim \text{CAR}(\rho, \sigma_u^2, W)$ con W matriz de vecindades
- Distribuciones iniciales : $B | \Sigma \sim \mathbf{N}_{m \times p}(B_0, D \otimes \Sigma)$, $\Sigma \sim \text{IW}(Q_0, m, \nu)$,
 $\sigma_v^2 \sim \text{IGa}(\alpha_0, \beta_0)$ y $\sigma_u^2 \sim \text{IGa}(\alpha_0, \beta_0)$ independientes

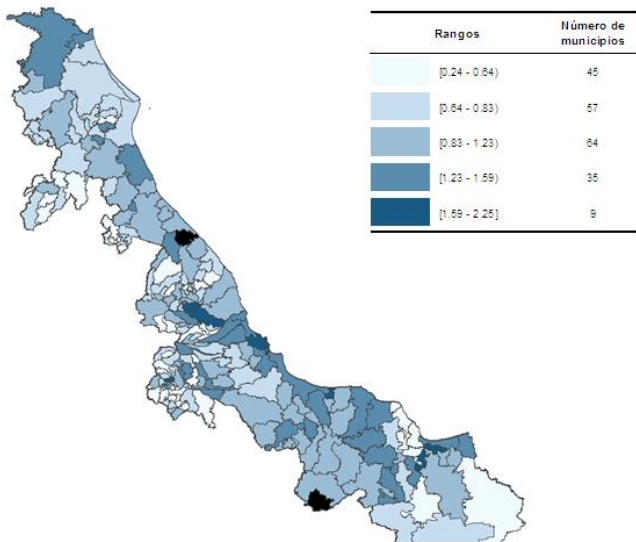
Carencias por rezago educativo



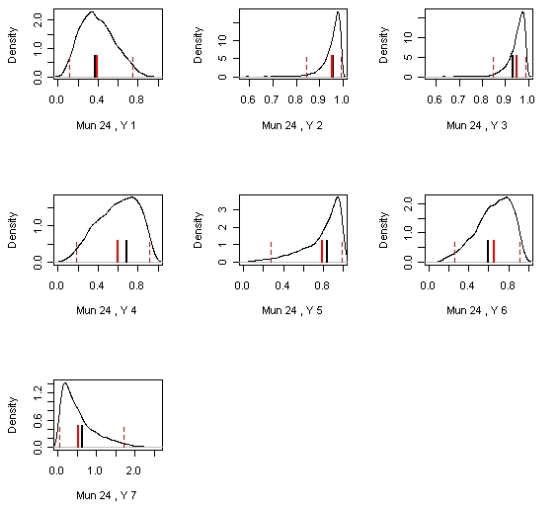
Carencias por alimentación



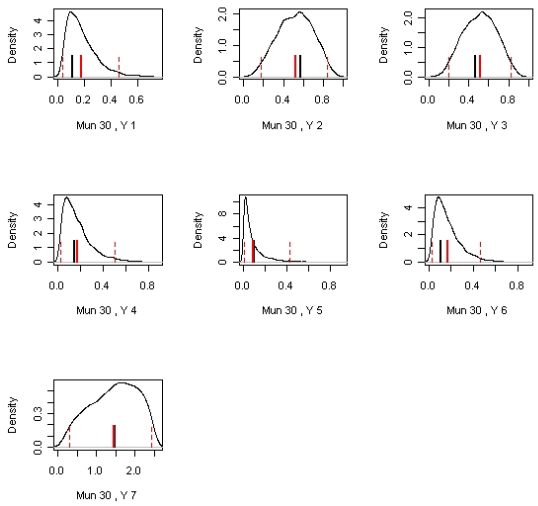
Ingreso corriente total



Estimaciones : Veracruz Municipio 24



Estimaciones : Veracruz Municipio 30



Idea

- De los 7 indicadores de pobreza, 4 se pueden obtener directamente del Censo ampliado
- El objetivo es estimar los indicadores de ingreso, seguridad social y acceso a alimentación
- Aunque los indicadores se reportan a nivel municipal, a veces se requiere desagregarlos por sexo, edad o pertenencia étnica
- Definir los 3 indicadores de pobreza $\mathbf{Y}_i = (Y_{1i}, Y_{2i}, Y_{3i})$ y un conjunto de variables explicativas $\mathbf{X}_i = (X_{1i}, X_{2i}, \dots, X_{pi})$ a 3 niveles (persona, hogar y municipal) para todos los individuos del MCS-ENIGH
- Definir las mismas variables explicativas \mathbf{X}_i que en el MCS-ENIGH, pero para todos los individuos del Censo ampliado y predecir \mathbf{Y}_i
- Calibrar los indicadores obtenidos del Censo ampliado usando los indicadores agregados del MCS-ENIGH a nivel estatal

Modelo estadístico

- Transformar : $\mathbf{Y}_i \rightarrow \mathbf{W}_i$, $W_{1i} = \log(Y_{1i} + b)$, $W_{ki} > 0 \iff Y_{ki} = 1$ and $W_{ki} \leq 0 \iff Y_{ki} = 0$, $k = 2, 3$

- Modelo multivariado :

$$\mathbf{W}_i | \theta_i, \Sigma \stackrel{\text{ind}}{\sim} N_m(\theta_i, c_i \Sigma), \quad i = 1, \dots, n$$

- Σ matriz de var-cov y c_i un factor de escala (diseño)

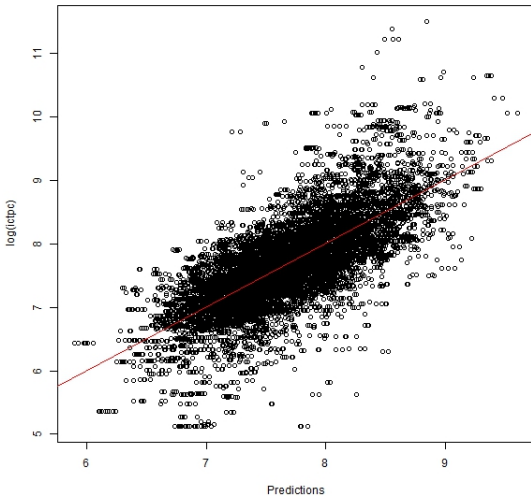
- Media

$$\theta_i = B\mathbf{x}_i + \mathbf{v}_{z_i}$$

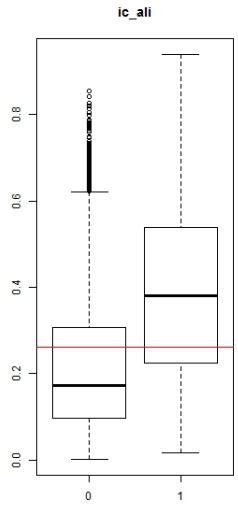
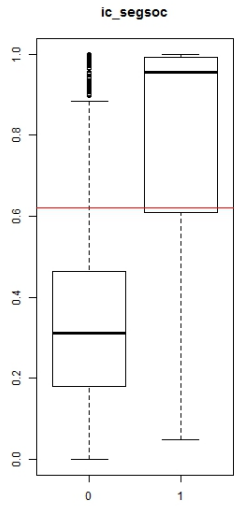
- Efectos fijos $B\mathbf{x}_i$ con B matriz de coeficientes
- Efectos estratificados $\mathbf{v}_j \stackrel{\text{iid}}{\sim} N_3(0, \Sigma_\nu)$, $j = 1, \dots, J$
- z_i = tamaño de localidad (rural, semi-urbana, urbana y metropolitana)

- Distribuciones iniciales : $B | \Sigma \sim N_{p \times 3}(B_0, D \otimes \Sigma)$, $\Sigma = (\sigma^2, \sigma_{12}, \sigma_{13}, \sigma_{23})$, $\sigma^2 \sim \text{Ga}(a_0, b_0)$, $\sigma_{12} \sim \text{Un}(-\sigma, \sigma)$, $\sigma_{13} \sim \text{Un}(-\sigma, \sigma)$, $\sigma_{23} \sim \text{Un}(-1, 1)$ and $\Sigma_\nu \sim \text{IW}(Q_\nu, 3, \omega_\nu)$

Ingreso predicho vs. observado



Probabilidades de éxito



Modelo municipal vs. individual

- Modelo a nivel municipal
 - Es más preciso
 - Tiempo de corrida es menor
 - No puede ser desagregado a otros subdominios
- Modelo a nivel individual
 - Tiene más incertidumbre (menos preciso)
 - Tiempo de corrida puede ser demasiado o incluso no operable
 - Puede ser agregado a varios dominios