



GOBERNACIÓN
SECRETARÍA DE GOBERNACIÓN



CONAPO
CONSEJO NACIONAL DE POBLACIÓN



Índice de marginación por localidad 2020

Nota técnico-metodológica

Septiembre de 2021

Introducción

El índice de marginación es una medida-resumen que permite diferenciar las localidades¹ del país según el impacto global de las carencias que padece la población como resultado de la falta de acceso a la educación, la residencia en viviendas inadecuadas y la carencia de bienes. Para su estimación se utilizó como fuente de información los microdatos del Censo de Población y Vivienda 2020.

Este nuevo ejercicio también utiliza el método de Distancia de Pena Trapero (1977), conocido como método DP_2 , el cual se usó para la estimación de los índices de marginación por entidad federativa, municipio y urbana 2020 (CONAPO, 2021a y 2021b). Este modelo estadístico mide distancias y permite comparar dos o más observaciones e interpretar sus diferencias cardinalmente. Al mismo tiempo, ofrece resultados similares a los ejercicios anteriores, estadísticamente es robusto y presenta múltiples bondades matemáticas. Con esto, se mantiene la naturaleza y utilidad del índice de marginación, y se atienden los principales aspectos de mejora técnica y metodológica de las estimaciones anteriores (Camberos y Bracamontes, 2007; Cárdenas, 2010; Gutiérrez-Pulido y Gama-Hernández, 2010).

El *índice de marginación por localidad 2020* es una herramienta de gran utilidad que permite observar y analizar la desigualdad en la participación del proceso de desarrollo y el disfrute de sus beneficios al interior de los municipios, así como su distribución territorial.

En esta nota técnico-metodológica se plantea, en primera instancia, una explicación sobre la construcción de los indicadores socioeconómicos que conforman el índice de marginación por localidad, así como una descripción de los métodos utilizados para su obtención, su estratificación –que determina los cinco grados de marginación– y su normalización. En un último apartado, se presentan los principales resultados a través de un cuadro que da cuenta del número de localidades y su monto de población según su grado de marginación, además de un mapa nacional que muestra la distribución territorial de las localidades según su grado de marginación.

Metodología de estimación del índice de marginación

El *Índice de marginación por localidad 2020* se desarrolló con indicadores socioeconómicos del Censo de Población y Vivienda 2020, obtenidos a través de una petición de información al Laboratorio de Microdatos del INEGI. Gracias a ello se estimaron los indicadores que representan ocho formas de exclusión de la marginación en las dimensiones: educación, vivienda y disponibilidad de bienes.

¹ La localidad se define como todo lugar ocupado con una o más viviendas, las cuales pueden estar habitadas o no; este lugar es reconocido por un nombre dado por la ley o la costumbre (INEGI, 2021).





Los ocho indicadores propuestos mantienen las bases teórico-conceptuales de los trabajos previos de marginación por localidad (CONAPO, 2002, 2007 y 2012); no obstante, en esta actualización algunos indicadores cambiaron su diseño de cálculo.

Para los ocho indicadores se estableció a I_j^i como el indicador socioeconómico j , para la localidad i , con $j = 1, 2, \dots, 8$ e $i = 1, 2, \dots, 189\ 041$ para el ejercicio del 2020. Debido al principio de confidencialidad estipulado en la Ley del Sistema Nacional de Información Estadística y Geográfica, en este análisis se estimaron los ocho indicadores simples y la marginación para dos tipos de localidades:

- a) Aquellas con más de tres viviendas particulares habitadas y para las cuales fue posible calcular sus indicadores de marginación de forma individual. Con lo cual se identificaron 108 144 localidades, que representan 57.1 por ciento de las localidades del listado de Integración Territorial (ITER) del censo 2020,² en donde se asienta 99.7 por ciento de la población total del país.
- b) Y, por otro lado, el conjunto de localidades con menos de tres viviendas particulares habitadas, las cuales suman 80 897 localidades que representan 42.7 por ciento del total de localidades registradas en el censo 2020. Estas localidades se agruparon a nivel municipal con una clave especial para incorporarlas al universo de estudio, obteniéndose 2 095 registros cuya población asciende a 391 046 habitantes que representan apenas 0.3 por ciento de la población total.

Con estos datos, el ejercicio contó con la información de 189 041 localidades, casi 99.8 por ciento de las localidades censadas; es decir, solo 391 localidades no formaron parte del universo de estudio porque no contaban con información de viviendas particulares.

A continuación, se presenta la forma del cálculo de los ocho indicadores usados para la estimación del *índice de marginación por localidad 2020*.

- Porcentaje de población analfabeta de 15 años o más (I_1^i)

El indicador se obtiene dividiendo el monto de población de 15 años o más que declaró no saber leer y escribir un recado, entre la diferencia de la población total de 15 años o más y aquellos que no especificaron su condición de alfabetismo:

$$I_1^i = \frac{PAN_{15+}^i}{P_{15+}^i - PNEALF_{15+}^i} \times 100$$

donde:

PAN_{15+}^i : es la población de 15 años o más analfabeta,
 P_{15+}^i : es la población total de 15 años o más, y

² El producto "Principales resultados por localidad "ITER" del censo 2020 está conformado por 189 432 localidades.





$PNEALF_{15+}^i$: es la población de 15 años o más que no especificó su condición de alfabetismo.

- Porcentaje de población de 15 años o más sin educación básica (I_2^i)

El indicador mide la magnitud de la población sin educación básica completa. Su cálculo se realiza en dos etapas. En la primera, la población que no especificó su último grado aprobado en secundaria o en estudios técnicos o comerciales con primaria terminada, se distribuye entre la población que aprobó entre uno y dos grados en estos mismos niveles educativos, aplicando la siguiente fórmula:

$$PSI_{15+}^i = PSCI_{15+}^i + \left[\frac{PSCI_{15+}^i}{PSCI_{15+}^i + PSCC_{15+}^i} \times PNEGS_{15+}^i \right],$$

donde:

PSI_{15+}^i : es la población de 15 años o más que aprobó entre el primer y segundo grado de secundaria o estudios técnicos o comerciales con primaria terminada con los no especificados de estos niveles educativos ya distribuidos,

$PSCI_{15+}^i$: es la población de 15 años o más que declaró haber aprobado entre el primer y segundo grado de secundaria o estudios técnicos o comerciales con primaria terminada,

$PSCC_{15+}^i$: es la población de 15 años o más que cursó el tercer grado en secundaria o tres o cuatro grados en estudios técnicos o comerciales con primaria terminada, y

$PNEGS_{15+}^i$: es la población de 15 años o más que no especificó su último grado cursado en secundaria o en estudios técnicos o comerciales con primaria terminada.

Con el dato de la población con estudios trancos en secundaria o en estudios técnicos o comerciales con primaria terminada, se procedió a calcular el indicador de porcentaje de población sin educación básica. Este porcentaje se calcula dividiendo la población de 15 años o más sin educación básica, entre la diferencia de la población total de 15 años o más y aquellos que no especificaron su nivel educativo:

$$I_2^i = \frac{PSIN_{15+}^i + PPI_{15+}^i + PSI_{15+}^i}{P_{15+}^i - PNEIN_{15+}^i} \times 100,$$

donde:

$PSIN_{15+}^i$: es la población de 15 años o más sin instrucción,

PPI_{15+}^i : es la población de 15 años o más con algún grado en educación primaria,

PSI_{15+}^i : es la población de 15 años o más con nivel incompleto de secundaria o estudios técnicos o comerciales con primaria terminada,

P_{15+}^i : es la población total de 15 años o más, y

$PNEIN_{15+}^i$: es la población de 15 años o más que no especificó su nivel de instrucción.





- Porcentaje de ocupantes en viviendas particulares sin drenaje ni sanitario (I_3^i)

Este porcentaje se obtiene al dividir el número de ocupantes de viviendas particulares sin drenaje ni sanitario, entre el número de ocupantes en viviendas particulares, menos el número de ocupantes de viviendas particulares en donde no se especificó la disponibilidad de drenaje ni sanitario:

$$I_3^i = \frac{OVSDS^i}{OVP^i - ONEDS^i} \times 100,$$

donde:

$OVSDS^i$: son los ocupantes de viviendas particulares sin disponibilidad de drenaje ni sanitario,

OVP^i : es el total de ocupantes en viviendas particulares, y

$ONEDS^i$: son los ocupantes de viviendas particulares en donde no se especificó la disponibilidad de drenaje ni sanitario.

- Porcentaje de ocupantes en viviendas particulares sin energía eléctrica (I_4^i)

Este indicador se obtiene al dividir el número de ocupantes de viviendas particulares sin disponibilidad de energía eléctrica, entre el número de ocupantes en viviendas particulares menos el número de ocupantes de viviendas particulares en donde no se especificó la existencia de luz eléctrica:

$$I_4^i = \frac{OSEE^i}{OVP^i - ONEEE^i} \times 100,$$

donde:

$OSEE^i$: son los ocupantes de viviendas particulares sin disponibilidad de energía eléctrica,

OVP^i : es el total de ocupantes en viviendas particulares, y

$ONEEE^i$: es el número de ocupantes de viviendas particulares en donde no se especificó la disponibilidad de energía eléctrica.

- Porcentaje de ocupantes en viviendas particulares sin agua entubada (I_5^i)

Para obtener este indicador se divide el número de ocupantes en viviendas particulares que no disponen de agua entubada, entre la diferencia del total de ocupantes en viviendas particulares y el total de ocupantes en viviendas en donde no se especificó la disponibilidad de agua entubada:

$$I_5^i = \frac{OSAE^i}{OVP^i - ONEAE^i} \times 100,$$

donde:





$OSAE^i$: son los ocupantes de viviendas particulares sin disponibilidad de agua entubada,
 OVP^i : es el total de ocupantes en viviendas particulares, y
 $ONEAE^i$: son los ocupantes de viviendas particulares en donde no se especificó la disponibilidad de agua entubada.

- Porcentaje de ocupantes en viviendas particulares con piso de tierra (I_6^i)

En este indicador se identifican los ocupantes de viviendas particulares con piso de tierra y se divide entre el total de ocupantes en viviendas particulares, menos el número de ocupantes de viviendas particulares en donde no se especificó el material predominante en pisos:

$$I_6^i = \frac{OPT^i}{OVP^i - ONEMP^i} \times 100,$$

donde:

OPT^i : son los ocupantes de viviendas particulares con piso de tierra,
 OVP^i : es el total de ocupantes en viviendas particulares, y
 $ONEMP^i$: son los ocupantes de viviendas particulares en las que no se especificó el material predominante en pisos.

- Porcentaje de ocupantes en viviendas particulares con hacinamiento (I_7^i)

El procedimiento para el cálculo de este indicador constó de dos etapas. Primero, en cada vivienda particular habitada se dividió el número de ocupantes entre el número de cuartos dormitorio, para identificar las viviendas con hacinamiento:

$$VHAC = \frac{OVP}{CDVP},$$

donde:

$VHAC$: es la vivienda con hacinamiento,
 OVP : es el número de ocupantes que residen habitualmente una vivienda particular, y
 $CDVP$: es el número de cuartos dormitorio en una vivienda particular.

En la segunda fase de cálculo, se dividió el número de ocupantes en viviendas particulares con hacinamiento, entre el total de ocupantes en viviendas particulares habitadas menos el número de ocupantes en viviendas particulares habitadas donde no se especificó el número de cuartos dormitorio:

$$I_7^i = \frac{OVHAC^i}{OVP^i - ONECD^i} \times 100,$$

donde:

$OVHAC^i$: son los ocupantes en viviendas particulares con hacinamiento,
 OVP^i : es el total de ocupantes en viviendas particulares, y





ONECDⁱ: son los ocupantes en viviendas particulares en donde no se especificó el número de cuartos dormitorio.

- Porcentaje de ocupantes en viviendas particulares sin refrigerador (I_8^i)

En este indicador se identifican los ocupantes en viviendas particulares que no disponen de refrigerador y se divide entre el total de ocupantes en viviendas particulares menos los ocupantes en viviendas particulares que no especificaron la disponibilidad de refrigerador:

$$I_8^i = \frac{OSR^i}{OVP^i - ONER^i} \times 100,$$

donde:

OSRⁱ: son los ocupantes de viviendas particulares sin disponibilidad de refrigerador,

OVPⁱ: es el total de ocupantes en viviendas particulares, y

ONERⁱ: son los ocupantes de viviendas particulares en donde no se especificó la disponibilidad de refrigerador.

Estimación del índice de marginación por localidad

Para estimar los índices de marginación en sus diferentes desagregaciones, se utilizó el Método de Distancia de Pena Trapero (1977), también conocido como método de medición de distancia P_2 o método DP_2 , desarrollado para hacer comparaciones intertemporales y espaciales y ha sido aplicado en investigaciones sobre calidad de vida en la Unión Europea y España (Somarriba y Pena 2009; Nayak y Mishra, 2012; Zarzosa, 1996 y 2012).

El método DP_2 es un modelo multivariado. El indicador construido con este procedimiento de agregación basado en un método econométrico de medición de distancias P_2 , posibilita la comparación en función de una amplia gama de variables en un indicador sintético: el indicador DP_2 . Por lo tanto, las unidades geográficas de un mismo nivel de desagregación se pueden contrastar en el tiempo por un conjunto completo de indicadores sintetizados en un índice.

La esencia de este método es medir la distancia entre cada área de estudio y una base de referencia. Como Zarzosa (2009) indica, la base de referencia puede conformarse por los “máximos valores observados para todas las variables o mayores valores que ellos; o bien, pueden ser los mínimos valores observados para todas las variables, o menores valores que ellos”. Incluso, la base de referencia puede ser un país o territorio ficticio con estas características. Si se opta por una base de referencia con las condiciones teóricas inferiores, entonces, el indicador DP_2 calcula la distancia de cada territorio estudiado a este territorio ficticio cuyo valor $DP_2 = 0$. Por lo tanto, un valor de DP_2 más alto indica mayor bienestar, dado que tiene una mayor distancia de la peor condición teórica.

Para permitir la comparación en el tiempo de los indicadores simples, se tomó como referencia la fecha censal 2010 y se determinó el vector base de referencia común en las condiciones ya conocidas de ser igual o inferior al mínimo, es decir, el peor escenario en las localidades (véase cuadro 1). De esta manera, se calcularon los valores de las distancias para





cada componente y para cada unidad geográfica. Es importante señalar que el cambio de base de referencia no condiciona o afecta la estructura del cálculo del índice de marginación, ya que cumple el postulado de invarianza respecto a la base de referencia.

Cuadro 1. Base de referencia a nivel localidad, 2010

Indicadores socioeconómicos	Base de referencia
Porcentaje de población analfabeta de 15 años o más	-100
Porcentaje de población de 15 años o más sin educación básica	-100
Porcentaje de ocupantes en viviendas sin drenaje ni excusado	-100
Porcentaje de ocupantes en viviendas sin energía eléctrica	-100
Porcentaje de ocupantes en viviendas sin agua entubada	-100
Porcentaje de ocupantes en viviendas con piso de tierra	-100
Porcentaje de ocupantes en viviendas con hacinamiento	-100
Porcentaje de ocupantes en viviendas sin refrigerador	-100

Fuente: Estimaciones del CONAPO con base en el INEGI, Censo de Población y Vivienda 2010.

El indicador sintético DP_2 se define de la siguiente manera:

$$DP_2 = \sum_{i=1}^n \frac{d_{ij}}{\sigma_j} (1 - R_{j,j-1,\dots,1}^2); \text{ con } R_1^2 = 0$$

donde:

$d_{ij} = |x_{rj} - x_{*j}|$: es la distancia de la j -ésima variable de la localidad r con respecto a la base de referencia $x_* = (x_{*1}, x_{*2}, \dots, x_{*n})$. Tomando como punto de referencia el valor mínimo de la variable, siendo esta la peor situación teórica,

σ_j : es la desviación estándar de la variable j ,

$R_{j,j-1,\dots,1}^2$: es el coeficiente de determinación de la regresión del indicador parcial j con respecto a los otros indicadores ($j-1, j-2, \dots, 1$). Esta expresión es parte de la varianza del indicador parcial I_j que se explica linealmente por el resto de los indicadores parciales,

$(1 - R_{j,j-1,\dots,1}^2)$: es el factor corrector que evita la duplicidad, al eliminar la información parcial de los indicadores ya contenidos en los indicadores precedentes, y

$R_1^2 = 0$: porque la primera componente aporta toda la información y al no existir un componente previo su ponderación es la unidad.³

El método DP_2 ofrece múltiples ventajas: permite realizar comparaciones espaciales y temporales; admite variables expresadas en diferentes unidades de medida; evita duplicidad de información y tiene imparcialidad en el esquema de ponderación. El indicador sintético de distancia DP_2 también presenta las siguientes propiedades matemáticas: existencia, determinación, monotonía, unicidad, invarianza frente a la base de referencia, homogeneidad, transitividad, exhaustividad, aditividad y neutralidad (Somarriba y Pena 2009; Zarzosa y Somarriba 2013). Estas bondades del método DP_2 brindan mayores beneficios que el Análisis de Componentes Principales (Somarriba y Pena, 2009).

³ Los símbolos y notaciones matemáticas de este apartado se presentan íntegramente como en Pena (1977) y Zarzosa (1996).





Otro de los beneficios del método DP_2 es que permite identificar la importancia de cada variable en el fenómeno estudiado, su aportación y el impacto de ellas de manera agregada. Las variables son analizadas con el coeficiente de discriminación de Ivanovic. Este coeficiente se considera una medida de bondad de los indicadores, ya que sirve para cuantificar el poder discriminante o la cantidad de información aportada por cada variable al indicador final. Se calcula mediante:

$$CD_i(P) = \frac{2}{m(m-1)} \sum_{j,l>j}^{k_j} m_{ji} m_{li} \left| \frac{x_{ji} - x_{li}}{\bar{x}_i} \right|, \quad i \in \{1, \dots, n\}$$

donde:

- m : el número de unidades de análisis j (localidad),
- m_{ji} : es el número de unidades de análisis j (localidad), en el conjunto P ,
- x_{ji} : es el valor de la variable x_i en la unidad de análisis j (localidad), y
- k_j : es el número de diferentes valores que x_i toma en el conjunto P .

Esta medida oscila entre 0 y 2, que son los dos extremos de los casos teóricos en cuanto al poder discriminatorio (Zarzosa, 1996). Si una variable tiene el mismo valor para todas las unidades de análisis, $CD=0$, la variable no tendrá ningún poder de discriminación; en cambio, si $CD=2$, la variable ejercerá un poder discriminatorio completo.

El método DP_2 también determina el impacto de cada indicador simple como parte del indicador sintético. Para hacer esto, se calcula la cantidad de información global de Ivanovic-Pena relativa individual o valor α_i (Zarzosa, 1996), definida como:

$$\alpha_i = \frac{CD_i(1 - R_{i,i-1,\dots,1}^2)}{\sum_{i=1}^n CD_i(1 - R_{i,i-1,\dots,1}^2)}$$

Esta medida presenta valores entre 0 y 1. Su cálculo fusiona la información única de cada indicador simple, el poder discriminatorio mediante el coeficiente de discriminación de Ivanovic y las medidas de la cantidad de información relativa (combinada) que aporta cada indicador simple individualmente. Cada valor resultante indica el aumento que se produce cuando se incorpora cada variable de forma ordenada en el indicador sintético DP_2 ; la suma de todos los valores de α_i es la unidad.

Con la obtención del índice de marginación a través del método DP_2 , los valores se clasifican en cinco categorías ordinales con el método de Dalenius y Hodges (1959), para obtener el grado de marginación. Este método forma estratos de manera que la varianza sea mínima al interior de cada estrato y máxima entre cada uno de ellos, es decir, son lo más homogéneos posibles. Este procedimiento utiliza la raíz de las frecuencias acumuladas para la construcción de los estratos, por lo que se lleva a cabo para la división de la población en el estrato L . Esta es una solución aproximada de Dalenius y Hodges (1959) a las ecuaciones de Dalenius (1950). De acuerdo con Gunning y Horgan (2004), el límite superior de cada estrato se determinó con la siguiente expresión:





$$Q = \frac{1}{L} \sum_{i=1}^J \sqrt{f_i}$$

Sea un conjunto de estratos determinados por su límite superior,

$$Q, 2Q, \dots, (L-1)Q, (L)Q.$$

donde:

J : es el número de clases dentro del grupo de la variable ordenada X ,
 $f_i \in (1, \dots, J)$: es la frecuencia en cada clase J , y
 L : es el número de estratos.

La eficiencia del método de la raíz de las frecuencias acumuladas depende principalmente del número de clases dentro del grupo de la variable ordenada. Sin embargo, no hay un procedimiento estándar sobre cómo elegir el mejor valor para el número de clases, siendo esto una limitante del método de Dalenius y Hodges. Para medir el efecto del número de clases en la varianza de cada estrato se recurrió a un método iterativo para obtener un criterio de agrupación óptimo.

Para establecer los límites de los estratos (b_1, \dots, b_L) que minimicen la varianza del estimador, se utiliza la asignación de Neyman para determinar el tamaño de muestra óptimo. Sea la varianza del estimador:

$$V(\bar{x}_{st}) = \sum_h \left(\frac{N_h}{N} \right)^2 \frac{S_h^2}{n_h}$$

donde:

S_h^2 : es la varianza poblacional en el estrato h ,
 n_h : es el tamaño de muestra en el estrato h utilizada por la asignación de Neyman, y
 N_h : es el total de elementos en el estrato h , sea $N = \sum_{h=1}^L N_h$.

Si se asume que la distribución dentro de cada estrato es aproximadamente uniforme, los límites se obtienen tomando intervalos iguales en la función de la raíz de las frecuencias acumuladas. Los límites se resuelven de manera iterativa:

$$\frac{S_h^2 + (b_h - \bar{X}_h)^2}{S_h} = \frac{S_{h+1}^2 + (b_h - \bar{X}_{h+1})^2}{S_{h+1}} \text{ para } h = 1, \dots, L-1$$

donde:

b_h : es el límite superior en el estrato h ,
 \bar{X}_h : es la media poblacional en el estrato h , y
 S_h^2 : es la varianza poblacional en el estrato h .

El requisito de precisión generalmente se establece cuando el coeficiente de variación sea igual a un nivel especificado entre uno y diez por ciento (Hidiroglou y Kozak, 2018). Este criterio, numéricamente, es comparable en términos del método de Dalenius y Hodges. Por ello, se desarrolló un algoritmo para determinar los límites de los estratos y los tamaños de





muestra que minimicen la función y cumplan con los criterios previamente descritos (véase cuadro 2).

Cuadro 2. Número óptimo de clases a nivel localidad, 2020

Número de clases	Error estándar	C.V.
23	0.347	0.017

Fuente: Estimaciones del CONAPO.

Por otro lado, para hacer cumplir la definición de los límites y el número óptimo de clases basados en la media aritmética, es importante la identificación de datos atípicos porque podrían ocasionar resultados poco adecuados. Para contrarrestar este problema, Hubert y Vandervieren (2007) proponen el método de caja, el cual toma en cuenta el grado de asimetría de un conjunto de datos, concluyendo que los límites con los que se debe trabajar (véase cuadro 3) están dados por las siguientes expresiones:

$$[Q_1 - 1.5 e^{-4MC} IQR ; Q_3 + 1.5 e^{3MC} IQR] \text{ para } MC \geq 0$$

$$[Q_1 - 1.5 e^{-3MC} IQR ; Q_3 + 1.5 e^{4MC} IQR] \text{ para } MC < 0$$

donde:

Q_1 y Q_3 : hacen referencia al primer y tercer cuartil, respectivamente; la diferencia entre estos dos valores da como resultado el valor de espacio intercuartil (IQR), y MC (*medcouple*): cuantifica el grado de asimetría de una muestra univariable $\{x_1, x_2, \dots, x_n\}$.

Cuadro 3. Límites para el cálculo de estratificación a nivel localidad, 2020

Valores atípicos	Rango	
	Mínimo	Límite
5 364	0.00	11.70

Fuente: Estimaciones del CONAPO.

El valor MC se encuentra definido por la ecuación:

$$MC = \text{med}_{x_i \leq Q_2 \leq x_j} h(x_i, x_j)$$

donde:

Q_2 : es la mediana de la muestra, y x_i y x_j : son elementos de la muestra.

La función h está definida por la ecuación:





$$h(x_i, x_j) = \frac{(x_j - Q_2) - (Q_2 - x_i)}{x_j - x_i}, \quad \forall x_i \neq x_j$$

Otra forma de apreciar el índice de marginación es normalizando sus valores para ver objetivamente la evolución de cada unidad territorial (Somarriba *et al.*, 2013). La normalización se realiza utilizando un cambio de escala conocido como normalización mínima-máxima. Con este procedimiento el índice de marginación se escala a valores relativos con un rango de entre cero y uno, lo cual permite su comparación numérica y le da una propiedad adicional al índice de marginación. Al mismo tiempo, la normalización determina el mismo sentido que el índice obtenido por el método DP_2 , donde los valores cercanos a cero implican mayor marginación.

El proceso de normalización consiste en:

$$DP_2^{normalizado} = \frac{DP_2^i - \min(DP_2)}{\max(DP_2) - \min(DP_2)}$$

donde:

DP_2^i : es el valor del índice de marginación de la localidad i ,

$\min(DP_2)$: es el valor mínimo o peor escenario que puede tomar el índice, y

$\max(DP_2)$: es el valor máximo u objetivo que puede tomar el índice.

Para este tipo de cálculo es necesario conocer los rangos objetivos, es decir, los puntos más extremos que puede tomar el índice en el año de observación. Como se señaló anteriormente, estos valores pueden ser empíricos, históricos o bien ideales, dependiendo del objetivo de la medición. De antemano, se sabe que cada indicador simple toma valores de cero a 100 y, además, el método DP_2 ya proporcionó el orden de entrada de las variables. Usando estos criterios se estiman los puntos focales extremos que puede tomar el índice de marginación. Se sabe que la peor situación es cuando la unidad de análisis toca todos y cada uno de los valores del vector base de referencia común; en tanto, el valor máximo sería cuando todos los indicadores simples cumplieron su objetivo y, por ende, cada uno toma el valor cero. Con base en lo anterior, las estimaciones determinaron los valores extremos que puede tomar el índice de marginación presentados en el cuadro 4.

Es importante aclarar que el índice de marginación normalizado se puede tomar como una guía ya que es sensible a los valores atípicos, lo que en cierta medida hace que se amplifiquen los datos normalizados ante los casos más extremos. Pero debido a las propiedades del método y que el rango de cada uno de los índices es estrecho y lineal, se podría esperar que la normalización no genere ruido en las colas de la distribución.

Con la normalización del índice se puede analizar de forma sencilla la posición que cada localidad ocupa de acuerdo con su intensidad de marginación estimada, su distancia





respecto a otras localidades o hacia los valores ideales o más críticos, así como su evolución y comportamiento en el tiempo.

Cuadro 4. República Mexicana. Escenarios extremos del índice de marginación por localidad, 2020

Escenario	% Población analfabeta de 15 años o más	% Población de 15 años o más sin educación básica	% Ocupantes en viviendas particulares sin drenaje ni excusado	% Ocupantes en viviendas particulares sin energía eléctrica	% Ocupantes en viviendas particulares sin agua entubada	% Ocupantes en viviendas particulares con piso de tierra	% Ocupantes en viviendas particulares con hacinamiento	% Ocupantes en viviendas particulares sin refrigerador	Índice de marginación
Mínimo	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Máximo	5.15	3.03	3.42	3.03	2.69	2.79	3.02	3.05	26.18 ^{a/}

^{a/} La suma de los porcentajes de los ocho indicadores da como resultado el índice de marginación (máximo y mínimo).
Fuente: Estimaciones del CONAPO con base en el INEGI, Censo de Población y Vivienda 2020.

Principales resultados

La estimación del índice de marginación para las localidades censales con tres o más viviendas en el año 2020, confirma que México se encuentra marcado por una profunda desigualdad en la participación del proceso de desarrollo y el disfrute de sus beneficios. Los datos del cuadro 5 muestran que 12 868 localidades tienen grado de marginación muy alto, donde residen 1.1 millones de personas; asimismo, 15 542 localidades tienen grado de marginación alto, donde viven 3.4 millones de personas. En el extremo opuesto, 23 880 localidades con grado de marginación muy bajo son lugar de residencia de casi 95.4 millones de personas; y 31 178 localidades con grado de marginación bajo alojan a 18.0 millones de personas. Por su parte, el número de asentamientos con grado de marginación medio asciende a 24 676 localidades, donde viven 7.6 millones de personas.

Cuadro 5. República Mexicana. Localidades y población por grado de marginación, 2020

Grado de marginación	Localidades		Población	
	Absolutos	Relativos	Absolutos	Relativos
Muy alto	12 868	11.9%	1 142 588	0.9%
Alto	15 542	14.4%	3 405 329	2.7%
Medio	24 676	22.8%	7 639 443	6.1%
Bajo	31 178	28.8%	18 043 587	14.4%
Muy bajo	23 880	22.1%	95 358 728	75.9%
Total	108 144	100%	125 589 675	100%

Fuente: Estimaciones del CONAPO con base en el INEGI, Censo de Población y Vivienda 2020.

El mapa 1 presenta la distribución territorial de las localidades según su grado de marginación. Es clara la incidencia de la marginación en las zonas montañosas de la Sierra

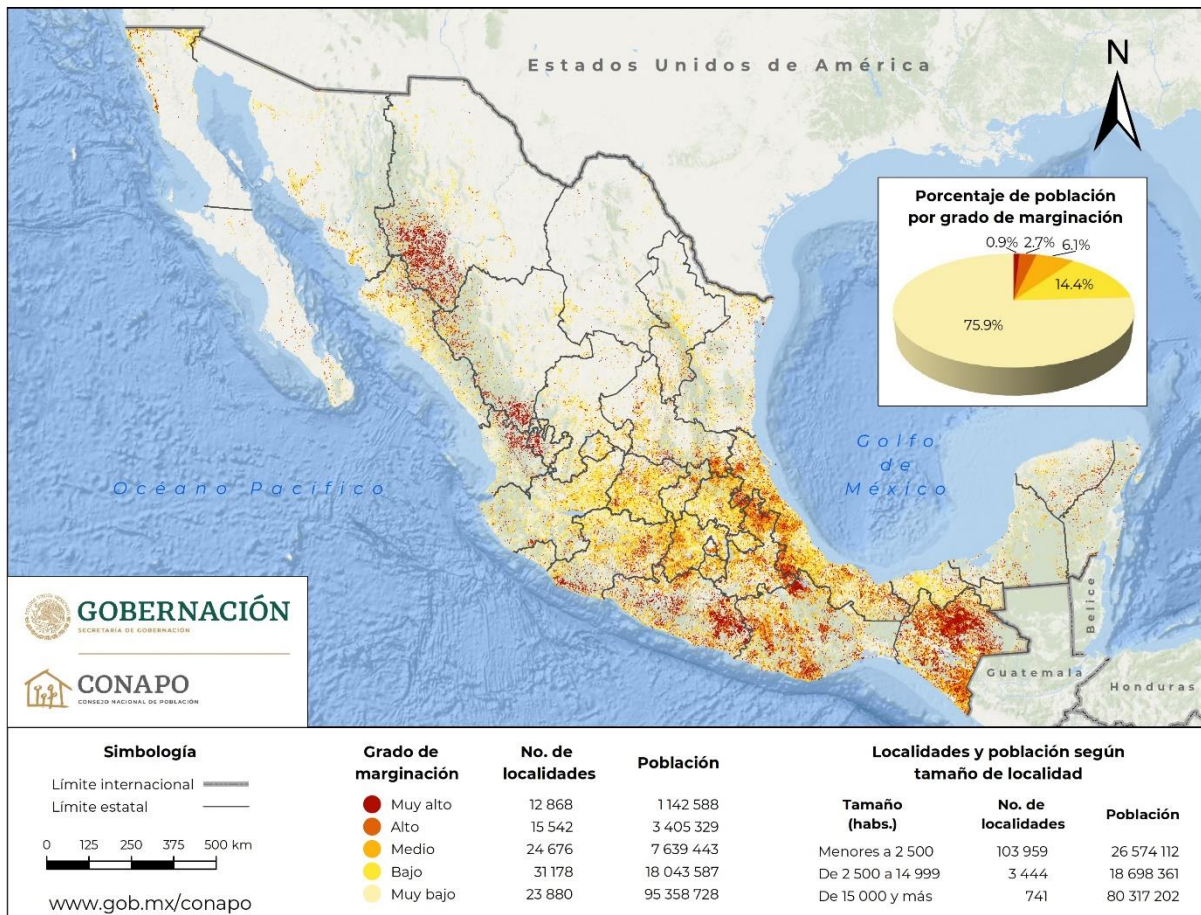




Madre Occidental en los estados de Chihuahua, Durango, Nayarit y Jalisco; en la Sierra Madre del Sur, destacando Guerrero y Oaxaca; en la Sierra Madre Oriental, principalmente en Hidalgo, Puebla y Veracruz; y de manera sobresaliente en algunas zonas del norte y centro de Chiapas.

En este ejercicio, cinco localidades se ubicaron en la peor situación de marginación teórica posible: Las Cruces, municipio Del Nayar, Nayarit; una localidad sin nombre, municipio Mezquitic, Jalisco; Cumbre de Guamúchil, municipio de Batopilas de Manuel Gómez Morín; Sicochi, municipio Guachochi; y Mesa de Marcelo, municipio Guadalupe y Calvo, estas tres localidades ubicadas en Chihuahua.

Mapa 1. República Mexicana. Grado de marginación por localidad, 2020



Fuente: Estimaciones del CONAPO con base en el INEGI, Censo de Población y Vivienda 2020.





Bibliografía y fuentes consultadas

- Camberos, M. y Bracamontes, J. (2007). Marginación y políticas de desarrollo social: un análisis regional para Sonora. *Problemas del Desarrollo*, 38 (149), 113-135. Recuperado de: http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0301-70362007000200006
- Cárdenas, O. (2010). Cardenalización del índice de marginación: una metodología para evaluar la eficiencia del gasto ejercido en el Ramo 33. *EconoQuantum*, 7 (1), 43-68. Recuperado de: <https://doi.org/https://doi.org/10.18381/eq.v7i1.119>
- CONAPO [Consejo Nacional de Población]. (2002). *Índice de marginación a nivel localidad, 2000*.
_____. (2007). *Índice de marginación a nivel localidad 2005*.
_____. (2012). *Índice de marginación por localidad 2010*.
_____. (2021a). *Índice de marginación por entidad federativa y municipio 2020* (en prensa).
_____. (2021b). *Índice de marginación urbana 2020* (en prensa).
- Dalenius, T. (1950). The problem of optimum stratification. *Scandinavian Actuarial J.*, 3-4, 203-13. Recuperado de: <https://doi.org/10.1080/03461238.1950.10432042>
- Dalenius y Hodges, J. L., Jr. (1959). Minimum variance stratification. *Journal of the American Statistical Association*, 54, 88-101.
- Gunning, P. y Horgan, J. M. (2004). A new algorithm for the construction of stratum boundaries in skewed populations. *Survey Methodology*, 30 (2), 159-166.
- Gutiérrez-Pulido, H. y Gama-Hernández, V. (2010). Limitantes de los índices de marginación de Conapo y propuesta para evaluar la marginación municipal en México. *Papeles de Población*, 16 (66), 227-257. Recuperado de: http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1405-74252010000400008
- Hidiroglou, M. A. y Kozak, M. (2018). Stratification of Skewed Populations: A Comparison of Optimization-based versus Approximate Methods. *International Statistical Review*, 86 (1), 87-105. Recuperado de: <https://doi.org/10.1111/insr.12230>
- Hubert, M. y Vandervieren, E. (2007). An adjusted boxplot for skewed distributions. *ScienceDirect*, 52 (12), 5186-5201. Recuperado de <https://doi.org/10.1016/j.csda.2007.11.008>
- INEGI [Instituto Nacional de Estadística y Geografía]. (2010). Censo de Población y Vivienda 2010.
_____. (2021). Censo de Población y Vivienda 2020.
- Nayak, P. y Mishra, S. K. (2012). Efficiency of Pena's P2 Distance in Construction of Human Development Indices. *SSRN Electronic Journal*. Recuperado de: <https://doi.org/10.2139/ssrn.2066567>
- Pena Trapero, J. B. (1977). *Problemas de la medición del bienestar y conceptos afines. Una aplicación al Caso Español*. Madrid: I. N. E.
- Somarriba, N. y Pena, B. (2009). Synthetic Indicators of Quality of Life in Europe. *Social Indicators Research*. Recuperado de: <https://doi.org/10.1007/s11205-008-9356-y>
_____, Zarzosa, P. y Pena, T. (2013). La calidad de vida en la Unión Europea. Un análisis temporal por medio de indicadores sintéticos. *Congreso de la Asociación Española de Ciencia Regional. XXXIX Reunión de Estudios Regionales. Smart regions for a smarter growth strategy: new challenges of Regional Policy and potentials of cities to overcome a worldwide economic crisis*. Recuperado de: <https://old.reunionesdeestudiosregionales.org/Oviedo2013/htdocs/pdf/p851.pdf>
- Zarzosa, P. (1996). *Aproximación a la medición del bienestar social*. Valladolid: Secretario de Publicaciones.
_____. (2009). Estimación de la pobreza en las comunidades autónomas españolas, mediante la distancia DP2 de Pena. *Estudios de Economía Aplicada*, 27 (2), 397-416.
_____. (2012). The Social Welfare in Spain before the Crisis: Territorial and Chronological Analysis. *International Journal of Advances in Management and Economics* 1 (4), 165-171.
_____ y Somarriba, N. (2013). An Assessment of Social Welfare in Spain: Territorial Analysis Using a Synthetic Welfare Indicator. *Social Indicators Research*, 111, 1-23.

