



**GOBERNACIÓN**  
SECRETARÍA DE GOBERNACIÓN



**CONAPO**  
CONSEJO NACIONAL DE POBLACIÓN



## **Índice de Intensidad Migratoria México-Estados Unidos por entidad federativa y municipio 2020**

### **Nota técnico-metodológica**

Septiembre de 2021

#### **Introducción**

La Secretaría General del Consejo Nacional de Población (SGCONAPO) elaboró por primera vez índices sintéticos desde los años noventa del siglo XX, tarea que continuó a principios del siglo XXI. Entre ellos se cuentan los índices de marginación, el índice de rezago reproductivo y el Índice de intensidad migratoria México-Estados Unidos. El impacto en materia de política pública de algunos de estos índices respondía a necesidades interinstitucionales, por lo que este fue inmediato, reflejándose en su uso extendido en distintas áreas de la Administración Pública Federal, como pieza importante para la instrumentación de acciones y diseño de estrategias específicas.

La investigación y análisis de la migración internacional México-Estados Unidos con interés de política pública es de vieja data, y los esfuerzos pioneros por sistematizar la información cuantitativa dieron cuenta del retorno, deportación y contratación masiva de migrantes de México desde los años de la revolución mexicana y hasta el usualmente llamado “Programa Bracero”. De manera más ordenada y continua, fue a finales de los años setenta del siglo pasado que —a través de encuestas en hogares y en puntos de reunión, cruce o devolución de personas migrantes—, se recabó información estadística sobre el fenómeno migratorio a la Unión Americana.

De la experiencia de la Encuesta Nacional de Emigración en la Frontera Norte de México (ENEFNEU), levantada por el Centro Nacional de Estadísticas del Trabajo (CENIET), a las encuestas *ad hoc*, como la Encuesta a Trabajadores Indocumentados Devueltos de Estados Unidos ETIDEU, levantada por la Secretaría de Gobernación (SEGOB), y hasta la incorporación de nuevas preguntas sobre el tema migratorio en el XI Censo de Población y Vivienda 1990 se logró una mejora sustancial en el conocimiento cuantitativo del proceso migratorio internacional desde el país.

La mejora en el conocimiento del fenómeno migratorio fue reforzada con el levantamiento de las Encuestas de la Dinámica Demográfica (ENADID) en 1992 y 1997 y el Conteo de Población y Vivienda de 1995 y la muestra que le acompañó, que se convirtieron en referentes para la investigación del proceso migratorio entre México y Estados Unidos. También en estos años, gracias al trabajo de El Colegio de la Frontera Norte (EL COLEF), a través del Proyecto Cañón Zapata y posteriormente junto con la Secretaría General del Consejo Nacional de Población y la Secretaría del Trabajo y Previsión Social, inició el levantamiento diario de la Encuesta sobre Migración en la Frontera Norte de México (EMIF), que debido a su continuidad se constituyó en una herramienta útil para estimar la dinámica migratoria en relación con los mercados de trabajo entre México y Estados Unidos.





Esta experiencia en el recabado del dato migratorio ayudó a desarrollar una mejora sustantiva en la comprensión de los distintos componentes del fenómeno migratorio, son los casos de la emigración hacia otros países —especialmente hacia Estados Unidos— y la circularidad que le acompaña (Bustamante, 1997), la inmigración internacional acumulada (población nacida en el extranjero) y reciente al país (cuyo uno de sus componentes es la migración de retorno de personas migrantes a México). Finalmente, estos diversos proyectos han permitido comprender también que México además de ser país de origen y retorno, es también territorio de tránsito y destino.

La riqueza de la información sobre el fenómeno migratorio en relación con las personas, los hogares y las viviendas como unidades de análisis permitió transitar del estudio de su magnitud hacia el de su intensidad a través de la selección de los indicadores más representativos del fenómeno migratorio. El análisis estadístico permitió tempranamente identificar que las correlaciones más fuertes se ubican entre las remesas, la emigración, la circularidad y el retorno migratorio con los Estados Unidos. Indicadores que hicieron factible la aplicación de técnicas estadísticas avanzadas (Análisis de Componentes Principales), de manera similar a como se había hecho en otros campos para definir un indicador sintético que diera cuenta integral del fenómeno.

En estricto sentido, la aplicación del Análisis de Componentes Principales para calcular el Índice de Intensidad Migratoria definió una ruta que permitió, con carácter exploratorio, dar cuenta de los elementos que modelan la estructura de la dinámica migratoria México-Estados Unidos. No obstante, a pesar de que la técnica recibió casi desde los inicios de su aplicación empírica diversas críticas (Mavrou, 2015) hasta la falta de comparabilidad del índice sintético en el tiempo (Hair et al., 1999), algunos de los puntos centrales a considerar en toda crítica a este método es cómo analizar la trayectoria de las cargas factoriales en el tiempo y la varianza común, aspectos técnica y metodológicamente sensibles.

Por ello, siguiendo la propuesta metodológica de los índices de marginación de la SG-CONAPO (2021) se decidió utilizar la técnica conocida como *Método de Medición de Distancia* o *método DP<sub>2</sub>*<sup>1</sup>, que permite obtener una medición cardinal para realizar comparaciones en el tiempo. Además, si se utiliza para medir las disparidades entre unidades político-administrativas similares, es obligatorio que la medida cuantifique tales disparidades. En este último sentido, afirma Zarzosa (1996) hay confianza en la medición cardinal que provee esta técnica, precisa “de forma aproximada”.

## Metodología

El Índice de Intensidad Migratoria México-Estados Unidos elaborado en la SG-CONAPO se deriva de estimaciones con base en los cuestionarios ampliados de los Censos de Población y Vivienda 2000, 2010 y 2020. Los hasta ahora publicados son los de 2000 con base en hogares y 2010 con base en viviendas censales, resultado del cambio conceptual y operativo introducido por el INEGI para el levantamiento de la información en ese año censal; y el Índice Absoluto de Migración Internacional 2010 que compara los indicadores básicos de 2000 y

<sup>1</sup> En adelante DP<sub>2</sub>.





2010. Conceptualmente, el Índice de Intensidad Migratoria México-Estados- Unidos es una medida sintética que informa integralmente del fenómeno migratorio entre ambos países, y con este carácter se constituyó en instrumento “puntual” de conocimiento de la dinámica de la migración internacional.

Debido a su propia construcción este permitió identificar entidades y municipios según su vínculo con la migración internacional desde México hacia Estados Unidos y de la Unión Americana hacia México, partiendo de la premisa de que las remesas proceden de ese país, y de la constatación empírica de que la emigración hacia el exterior es prácticamente de destino único, así como la continuidad de la circularidad migratoria entre ambos países, y el retorno de personas nacidas en México desde la Unión Americana. El análisis de las estimaciones de los tres cuestionarios ampliados corrobora la preminencia de dicha relación migratoria. Sin embargo, también permite advertir algunos cambios en la migración internacional y sus componentes de dinámica, como el creciente número de personas nacidas en el extranjero residentes en México, que ayudan a formar un panorama lo más completo posible del fenómeno migratorio.

El volumen estimado total de migrantes internacionales ha descendido de 1.62 millones a 1.11 millones y a 803 mil personas en los quinquenios 1995-2000, 2005-2010 y 2015-2020, respectivamente; vinculado a este descenso, el volumen de emigrantes internacionales también ha descendido mientras su participación porcentual en el volumen total de migrantes pasó de 81.7 por ciento a 65.2 y 70.3 por ciento en los quinquenios indicados. La estimación de población migrante circular, es decir, personas que en cualquier momento durante el quinquenio previo al censo migraron pero que a la fecha del evento censal residían en México, ha fluctuado de más a menos en cada uno de los periodos indicados, pasando de 17.4 por ciento, a 31.5 y 22.2 por ciento como porcentaje también del volumen total de personas migrantes. En relación con la migración de retorno —personas nacidas en México que cinco años antes del evento censal residían en el extranjero, representaron 0.3 por ciento, 0.8 y nuevamente 0.3 por ciento de la población de cinco años o más residente en el país (281 mil, 859 y 333 mil), en 2000, 2010 y 2020.

Específicamente, el volumen de emigrantes internacionales con destino a Estados Unidos y que residían en ese país al momento del levantamiento censal representó 97.9 por ciento, 94.5 y 85.1 por ciento del volumen estimado del total de emigrantes en cada quinquenio, respectivamente. En el caso de la población migrante circular procedente de Estados Unidos, los valores porcentuales estimados en cada quinquenio respecto a la población migrante circular total son 91.5 por ciento, 87.8 y 71.2 por ciento. Por último, la población migrante de retorno de Estados Unidos en 2000, 2010 y 2020 ha mantenido su estructura porcentual, representando 95.0, 95.9 y 88.6 por ciento del total de retornos con volúmenes estimados de 267 mil, 824 mil y 295 mil, respectivamente.

Explícitamente se identifica un cambio cuantitativo en la migración internacional desde y hacia México en conjunto y, en particular, en la relación migratoria con Estados Unidos. Por ello, antes de la construcción del Índice de Intensidad Migratoria, se analizó el cambio en los indicadores que se utilizan como base para la estimación de los vectores a nivel de vivienda. De estos cambios destaca que, de ser casi una migración de destino único, actualmente (2020) 15 de cada 100 emigrantes de México residen en un país distinto a Estados Unidos, aproximadamente tres de cada diez presentan un movimiento circular con





un país distinto a EE. UU. y uno de cada diez retorna de un país distinto de Estados Unidos. No obstante, las correlaciones más importantes son las referidas a la dinámica migratoria entre ambos países, motivo por el cual los indicadores se anclan a este resultado.

El presente desarrollo metodológico conserva los cuatro indicadores simples que integran el índice, cabe acotar que, técnicamente, todas las estimaciones inicialmente tienen como unidad de análisis a las personas, pero dado que el objetivo es aproximarse a la intensidad del fenómeno y no a su magnitud, estas se agregan a nivel de vivienda.

En relación con el cálculo de los cuatro indicadores para 2020, al igual que en 2000 y 2010, se estableció a  $I_{ij}$  como el indicador de migración internacional  $j$ , para la entidad o municipio  $i$ , con  $j=1, 2, \dots, 4$  e  $i=1, 2, \dots, 32$  (para entidades) o  $i=1, 2, \dots, 2469$  (para municipios).

#### Porcentaje viviendas con remesas internacionales ( $I_{i1}$ )

$$I_{i1} = \left( \frac{V_i^{remesas}}{V_i^{total} - V_{i1}^{s.i.c.}} \right) \times 100$$

donde:

$V_i^{remesas}$ : son las viviendas que reciben remesas desde otro país,

$V_i^{total}$ : es el total de viviendas en la unidad político-administrativa  $i$ ,

$V_{i1}^{s.i.c.}$ : es el número de viviendas en la unidad político-administrativa  $i$ , en las que no se dispone de información completa para este indicador.

#### Porcentaje viviendas con emigrantes internacionales del quinquenio anterior en Estados Unidos ( $I_{i2}$ )

$$I_{i2} = \left( \frac{V_i^{emigi}}{V_i^{total} - V_{i2}^{s.i.c.}} \right) \times 100$$

donde:

$V_i^{emigi}$ : son las viviendas en las que hubo emigrantes internacionales en el quinquenio anterior, personas que continúan residiendo en Estados Unidos,

$V_i^{total}$ : es el total de viviendas en la unidad político-administrativa  $i$ ,

$V_{i2}^{s.i.c.}$ : es el número de viviendas en la unidad político-administrativa  $i$ , en las que no se dispone de información completa para este indicador.





### Porcentaje viviendas con migrantes circulares de Estados Unidos del quinquenio anterior ( $I_{i3}$ )

$$I_{i3} = \left( \frac{V_i^{mig.cir}}{V_i^{total} - V_{i3}^{s.i.c.}} \right) \times 100$$

donde:

$V_i^{mig.cir}$ : son las viviendas con personas que emigraron a Estados Unidos y regresaron a residir a México en el quinquenio de referencia 2015-2020,

$V_i^{total}$ : es el total de viviendas en la unidad político-administrativa  $i$ ,

$V_{i3}^{s.i.c.}$ : es el número de viviendas en la unidad político-administrativa  $i$ , en las que no se dispone de información completa para este indicador.

### Porcentaje viviendas con migrantes de retorno de Estados Unidos ( $I_{i4}$ )

$$I_{i4} = \left( \frac{V_i^{mig.ret(eua)}}{V_i^{total} - V_{i4}^{s.i.c.}} \right) \times 100$$

donde:

$V_i^{mig.ret}$ : son las viviendas con personas nacidas en nuestro país que en marzo de 2015 vivían en otro país y para marzo de 2020 residían en México,

$V_i^{total}$ : es el total de viviendas en la unidad político-administrativa  $i$ ,

$V_{i4}^{s.i.c.}$ : es el número de viviendas en la unidad político-administrativa  $i$ , en las que no se dispone de información completa para este indicador.

### Estimación del índice de Intensidad Migratoria

Para estimar los índices de intensidad migratoria a nivel estatal y municipal se utilizó el Método de Distancia de Bernardo Pena Trapero (1977), también conocido como método de medición de distancia  $P_2$  o método  $DP_2$ , desarrollado para hacer comparaciones intertemporales, aplicado en investigaciones sobre calidad de vida (Zarzosa, 1996; Somarriba y Pena 2009; Nayak y Mishra, 2013; Marquina, et al., 2015) Recientemente la SGCONAPO (2021) utilizó esta técnica para calcular la marginación por entidad federativa y municipio como alternativa al método de componentes principales usado para calcular la familia de índices de marginación 2015 y anteriores.

El método  $DP_2$  forma parte de los métodos de análisis multivariantes de datos. El índice construido con este procedimiento de agregación basado en un método econométrico de medición de distancias  $P_2$ , posibilita la comparación en el tiempo en función de las variables que conceptualmente se han identificado deben formar parte de un



indicador sintético: el indicador  $DP_2$ . Por lo tanto, las unidades se pueden contrastar entre diferentes años por un conjunto completo de indicadores simples sintetizados en un índice.

La esencia de este método es medir la distancia entre cada unidad político-administrativa de estudio y una base de referencia. Zarzosa (1996) indica que en la primera aplicación de la  $DP_2$  se eligió como vector base o de referencia el de mínimos, que representa la situación "teórica menos deseada". Más tarde Zarzosa (2009) ha precisado que la base de referencia puede conformarse por los "máximos valores observados para todas las variables o mayores valores que ellos en el tiempo o en alguna unidad espacial similar (político-administrativa) de comparación; o bien, pueden ser los mínimos valores observados para todas las variables, o menores valores que ellos. Incluso, la base de referencia puede ser una unidad territorial inexistente o ficticia 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 que se tiene una mayor distancia a dicha condición teórica inferior. En sentido inverso, si la base de referencia son los máximos, un valor  $DP_2$  más alto indica un alejamiento de la condición teórica inferior.

Marquina *et al.* (2015) consideran que las decisiones adoptadas a la hora de definir las medidas sintéticas, como la elección del procedimiento de normalización en componentes principales o la determinación del vector de referencia en  $DP_2$ , deben evaluarse en función de la comparabilidad de los resultados obtenidos por las distintas técnicas multivariantes.

Para permitir la comparación en el tiempo de los indicadores simples, se tomó como referencia la fecha censal 2000 y se determinó como vector base de referencia común, en las condiciones ya conocidas de la migración internacional, los valores máximos, asimilándolos a una situación de máxima intensidad migratoria, si se toma en cuenta que el método  $DP_2$  inicia con la "Distancia de Frechet" (véanse cuadros 1 y 2).

**Cuadro 1. Vector de referencia a nivel entidad federativa, 2000**

Indicadores simples de migración internacional	Porcentaje
Porcentaje de viviendas que perciben remesas	13.51
Porcentaje de viviendas con emigrantes en Estados Unidos	11.73
Porcentaje de viviendas migrantes circulares de Estados Unidos	3.41
Porcentaje de viviendas con migrantes de retorno del quinquenio de Estados Unidos	2.62

Nota: Los valores estimados pueden no coincidir con los publicados a nivel de centésimas o milésimas debido a la estandarización para la obtención de los vectores en las tres muestras censales, 2000, 2010 y 2020.

Fuente: Estimaciones del CONAPO con base en el INEGI, Censo de Población y Vivienda 2000.





**Cuadro 2. Vector de referencia a nivel municipal, 2000**

Indicadores simples de migración internacional	Porcentaje
Porcentaje de viviendas que perciben remesas	54.49
Porcentaje de viviendas con emigrantes en Estados Unidos	48.68
Porcentaje de viviendas con migrantes circulares de Estados Unidos	34.72
Porcentaje de viviendas con migrantes de retorno del quinquenio de Estados Unidos	18.01

Nota: Los valores estimados pueden no coincidir con los publicados a nivel de centésimas o milésimas debido a la estandarización para la obtención de los vectores en las tres muestras censales, 2000, 2010 y 2020.

Fuente: Estimaciones del CONAPO con base en el INEGI, Censo de Población y Vivienda 2000.

El indicador de  $DP_2$  sintético 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 entidad federativa o municipio  $r$  con respecto a la base de referencia  $_*_*_* = (x_{*1}, x_{*2}, \dots, x_{*n})$ . Tomando como punto de referencia una situación teórica no deseable,

$\sigma_j$ , 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  $(j-1, -2, \dots, 1)$ . 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 pretende eliminar la doble información o redundancias, elimina 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.

Zarzosa (1996), Somarriba y Pena (2009), así como Zarzosa y Somarriba (2013) indican que el método  $DP_2$  permite realizar comparaciones temporales; admite variables expresadas en diferentes unidades de medida; evita la duplicación de información y tiene imparcialidad en el esquema de ponderación. De acuerdo con los autores del método, teóricamente y en su aplicación concreta, se demuestra que el indicador sintético tiene las siguientes propiedades matemáticas: existencia, determinación, monotonía, unicidad, invariancia





frente a la base de referencia, homogeneidad, transitividad, exhaustividad, aditividad y neutralidad.

En relación con estas propiedades el de neutralidad indica que en él método no se determina de forma arbitraria ni el peso de los indicadores parciales ni su ordenación, ya que en esta última interviene el valor absoluto del coeficiente de correlación lineal. Por otra parte, la invariancia señala que el indicador es invariante ante un cambio en la escala con que se miden los componentes del indicador sintético o que es invariante ante un cambio de la base de referencia de los componentes del indicador, que de acuerdo con Trapero (2009) debería ser de utilidad para analizar la bondad del método de agregación. La literatura especializada sugiere que como parte de los métodos multivariantes  $DP_2$  se complementa con la aplicación de otras técnicas.

Además, Zarzosa (1996) indica que el método  $DP_2$  permite identificar la importancia de cada variable en el fenómeno estudiado, su aportación y el impacto de ellas de manera agregada, permitiendo incluso identificar las que no resultan relevantes en relación con el fenómeno estudiado. Mediante el Coeficiente de Discriminación de Ivanovic se cuantifica el poder discriminante o la cantidad de información que aporta cada variable  $i$  al indicador final. Se expresa como:

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

donde:

$m$  es el número de unidades de análisis  $j$  (entidad o municipio),

$m_{ji}$  es el número de unidades de análisis  $j$  (entidad o municipio), en el conjunto  $P$ ,

$x_{ji}$  es el valor de la variable  $x_i$  en la unidad de análisis  $j$  (entidad o municipio), 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. En relación con el cálculo del IIM- $DP_2$  ninguna de las cuatro variables iniciales fue discriminada.

El método  $DP_2$  también permite determinar el impacto de cada indicador simple como parte del indicador sintético. Para hacer esto, se calcula, en el conjunto total ordenado de variables, la Cantidad de Información Global de Ivanovic-Pena Relativa Individual ( $\alpha$ ), que relaciona el coeficiente de discriminación  $CD_i$  y el factor corrector, 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 ( $CD_i$ ) y las medidas de la cantidad de información relativa (combinada) que aporta cada indicador simple individualmente (Cantidad de Información Global de Ivanovic-Pena). 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  $\alpha_i$  es la unidad.

Cabe acotar que en relación con las comparaciones entre unidades temporales Zarzosa (2009) ha comentado que satisfechas ciertas condiciones “el valor del indicador sintético de distancia entre dos unidades es igual a la diferencia, en valor absoluto, entre los valores de los indicadores sintéticos de distancia de las dos unidades calculados con respecto a la misma base de referencia” (p. 408).

En continuidad con la usual clasificación del índice de intensidad migratoria, una vez calculado el Índice  $DP_2$ , para obtener el grado de intensidad migratoria se procedió a conformar 5 estratos usando el método de Dalenius y Hodges (1959). Usualmente se indica que una de las limitaciones de este método es la inexistencia de un procedimiento estándar sobre cómo elegir el mejor valor para el número de clases en el rango de la variable ordenada, condición de lo cual depende principalmente la eficiencia del método de la raíz de las frecuencias acumuladas. Y aunque existen desarrollos que buscan solucionar esta limitante (Hoaglin, 1983; Machado et al., 2010) se recurrió a la asignación óptima de Neyman, bajo la cual la varianza es la mínima para cada estrato.

Los componentes para los cálculos de los valores límites de los estratos para el indicador sintético  $IIM - DP_2$  son los siguientes: la variable para estratificar, el número de estratos ( $L$ ), el número de clases ( $J$ ) y requisito de precisión estadística expresado en términos de un coeficiente de variación objetivo ( $CV$ ), que de acuerdo con Hidiroglou y Kozak (2018) se establece entre 0 y 10 por ciento, indicando una baja dispersión en la distribución esperada del indicador.

En el caso del  $IIM - DP_2$  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. El número de clases se eligió donde mediante este método se minimizará el coeficiente de variación para dado número de estratos en que se subdivide a una población  $N, N = N_1 + N_2 + N_3 + \dots + N_L$ .

La selección del tamaño de muestra de cada estrato se lleva a cabo bajo la asignación de Neyman, donde la asignación del tamaño de la muestra que minimiza la varianza de un estimador está dado por (Wright, 2014):

$$n_h = n * \frac{N_h * S_h}{\sum_{h=1}^L N_h * S_h}$$

donde:

$n$  es el total de la muestra,

$n_h$  es el tamaño de la muestra de un estrato  $h$ ,

$N_h$  es el tamaño de la población de un estrato  $h$ ,





$S_h$  es la desviación estándar del estrato  $h$ .

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. De modo que la implementación de la asignación de Neyman se lleva a cabo con  $q_1 = q_3 = 0.5$  y  $q_2 = 0$  como parámetros para la asignación óptima en cada estrato, siendo la varianza del estimador:

$$V(\bar{X}_{st}) = \sum_{h=1}^L \left(\frac{N_h}{N}\right)^2 \frac{S_h^2}{n_h}$$

Si se asume que la distribución dentro de cada estrato se distribuye aproximadamente de manera 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:

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

donde:

$b_h$ , es el límite superior en el estrato  $h$ ,

$\bar{X}_h$ , es la media poblacional en el estrato  $h$ ,

Obtenido el número de clases, se pasa a formar intervalos donde se distinguen las frecuencias de cada clase. De las frecuencias de cada clase ( $f_i$ ) se obtiene su raíz cuadrada y se acumula la raíz cuadrada de la frecuencia de cada clase a la cual se le llama  $C$ ,

$$C = \sum_{i=1}^J \sqrt{f_i} \quad (i = 1, \dots, J)$$

Para obtener los límites de cada estrato se divide  $C$  entre el número de estratos,

$$Q = C/L$$

donde  $L$  es el número de estratos y para obtener los límites de cada uno de estos estratos se sigue,



$$Q,$$

$$2 * Q \dots,$$

$$(L - 1) * Q,$$

$$L * Q$$

<b>Cuadro 3. Número óptimo de clases a nivel estatal, 2020</b>		
Número de clases	Error estándar	CV
10	0.1320157	0.006619189
Fuente: Estimaciones de Conapo con base en Conapo (2021) y Baillargeon, S. and Rivest, L. P. (2011).		

<b>Cuadro 4. Número óptimo de clases a nivel municipal, 2020</b>		
Número de clases	Error estándar	CV
75	0.2153676	0.003411577
Fuente: Estimaciones de Conapo con base en Conapo (2021) y Baillargeon, S. and Rivest, L. P. (2011).		

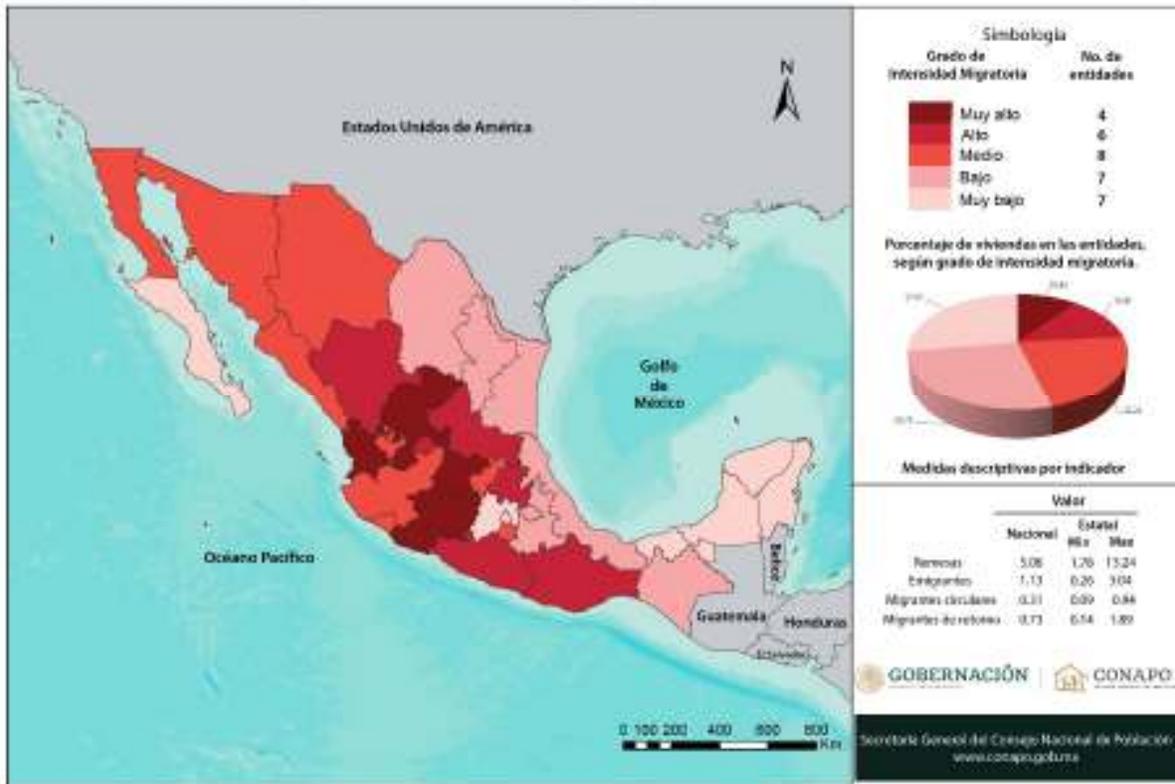
### **Índice de Intensidad Migratoria México-Estados Unidos por entidad federativa**

La estimación del índice de intensidad migratoria por entidad federativa señala que Zacatecas, Nayarit, Michoacán y Guanajuato se clasifican en 2020 como de “Muy alta” intensidad migratoria, las cuatro están ubicadas en lo que usualmente los estudios de migración internacional denominan región histórica o tradicional (Durand, 1998; Corona, 2000; CONAPO, 2004). De acuerdo con las proyecciones de población (2016-2050), en términos demográficos, estas cuatro entidades concentran once por ciento de la población total del país con una edad media situada por debajo de la media nacional (28.45 años). En el extremo opuesto con “Muy baja” intensidad migratoria se ubican Tabasco, Quintana Roo, Ciudad de México, Campeche, Yucatán, Baja California Sur y el estado de México, cuatro de ellas ubicadas en la región migratoria Sur-Sureste (véase mapa A1). Estas últimas, en general no se caracterizan por tener grandes flujos migratorios internacionales o una gran captación de remesas familiares, en su momento fueron categorizadas como regiones emergentes de migración internacional. E igualmente las proyecciones de población indican que en ellas reside alrededor de seis por ciento de la población del país. La edad media de la población en Yucatán (29.11 años) rebasa la media nacional y, Campeche y Tabasco, ambas con una media de 28 años, están próximas al promedio nacional.





Mapa A1. Grado de intensidad migratoria por entidad federativa, 2020



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

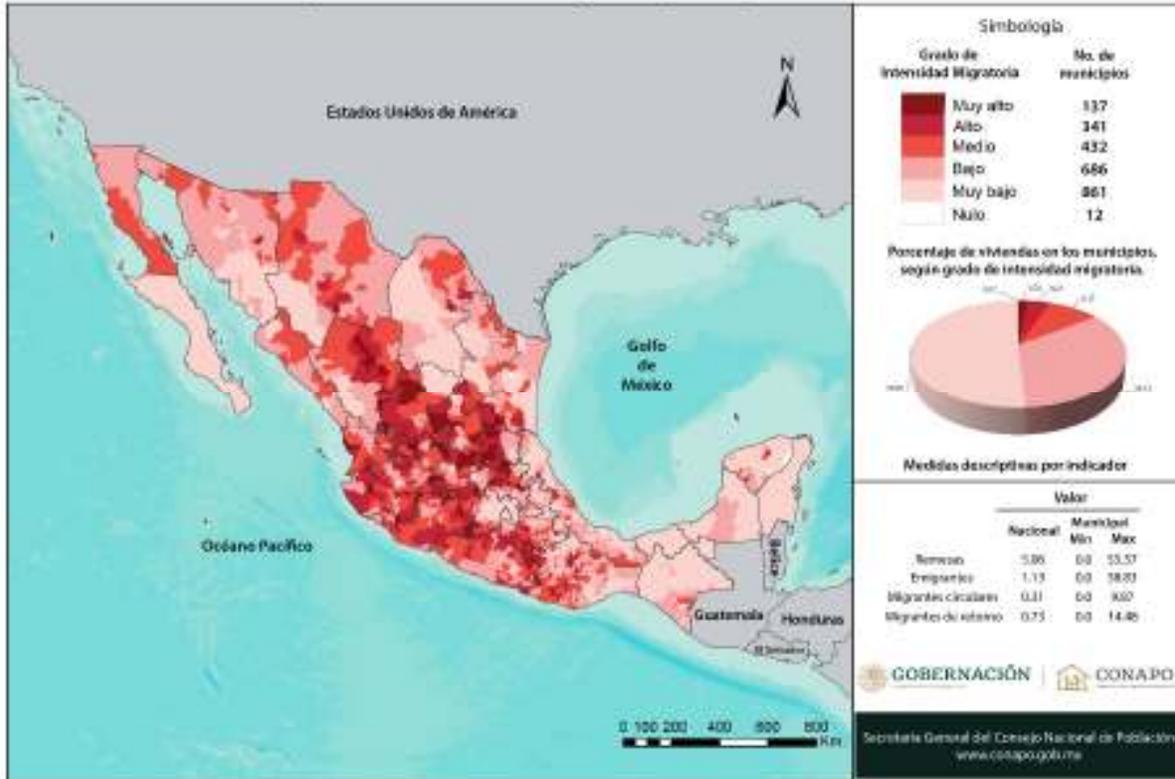
### Índice de intensidad migratoria por municipio

En general el fenómeno migratorio en niveles político-administrativos de menor tamaño muestra la relevancia de algunos territorios. En el caso de 2020, destaca que en similitud con 2010 (11) y a diferencia de 2000 (96) solo 12 municipios no presentaron valores estimados en ninguna de las variables consideradas para calcular el índice de intensidad migratoria, se trata de cinco municipios ubicados en Oaxaca (San Bartolo Yautepec, San Pedro Ocopetatlillo, Santa María Ixcatlán, Santa María Tataltepec y Santiago Tepetlapa), dos en Yucatán (Chemax y Chikindzonot), dos más en Chiapas (Totolapa y Honduras de la Sierra), uno en Campeche (Seybaplaya), uno en Tlaxcala (San Nicolás) y uno más en Tamaulipas (La Magdalena Tlaltelulco). En estricto sentido, se puede decir que, el fenómeno migratorio ha continuado en el periodo 2015-2020 abarcando prácticamente todo el territorio nacional. En el caso de la intensidad migratoria “Muy alta” ocho de cada diez de los 137 municipios clasificados con este grado se ubican en Oaxaca, Michoacán, Zacatecas, Hidalgo, San Luis Potosí, Guanajuato y Jalisco. En cambio, los municipios con grado “Muy bajo” (34.9% del total), se ubican principalmente en Oaxaca, Veracruz, el estado de México, Chiapas y Yucatán, que concentran poco más de dos terceras partes (64.2%) del total de los municipios en este grado (véase mapa A2).





Mapa A2. Grado de intensidad migratoria por municipio, 2020



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





## Referencias

Baillargeon, S. y Rivest, L.-P. (2011). The construction of stratified designs in R with the package stratification. Statistics Canada, Catalogue No. 12-001-X. Disponible en <https://www.mat.ulaval.ca/fileadmin/mat/documents/lrivest/Publications/102Bail-Rivest11.pdf>

Bustamante, J. A. (1997). El marco teórico--metodológico de la "circularidad migratoria": su validación empírica. *Sociológica: Revista de pensamiento social*, 2, 77-120. Recuperado de <http://hdl.handle.net/2183/2703>

\_\_\_\_\_. (2017). Migración irregular de México a Estados Unidos. Diez años de investigación del Proyecto Cañón Zapata. *Frontera Norte*, 12(23), 7-49. <https://doi.org/10.17428/rfn.v12i23.1393>

CONAPO [Secretaría General del Consejo Nacional de Población] (2002). Índice de Intensidad Migratoria México-Estados Unidos 2000.

\_\_\_\_\_. (2004). *La nueva era de las migraciones*. Ciudad de México.

\_\_\_\_\_. (2012). Índice de Intensidad Migratoria México-Estados Unidos 2010.

\_\_\_\_\_. (2014). Índice Absoluto de Intensidad Migratoria México-Estados Unidos 2010.

\_\_\_\_\_. (2021). Índice de marginación por entidad federativa y municipio 2020. Nota técnico-metodológica. Recuperado de <https://www.gob.mx/conapo/documentos/indice-de-marginacion-por-entidad-federativa-y-municipio-2020>

Corona, R. y Tuirán, R. (2000). Dimensión de la migración a Estados Unidos desde la perspectiva de los hogares. En: *Migración-México-Estados Unidos: Continuidad y cambio*. Ciudad de México.

Dalenius, T. (1950). The problem of optimum stratification. *Scandinavian Actuarial Journal*, 3-4, 203-13. Disponible en <https://doi.org/10.1080/03461238.1950.10432042>

Durand, J. (1998). ¿Nuevas regiones migratorias? En: René Zenteno (Coord.). *Población, desarrollo y globalización*. Tijuana, B.C.: SOMEDE-El Colegio de la Frontera Norte.

Hair, J. F. et al. (1999). *Análisis multivariante*. Madrid: Pearson-Prentice Hall.

Hoaglin, D. C., Mosteller, F. y Tukey, J. W. (1983). *Understanding Robust and Exploratory Data Analysis*. Ney York: John Wiley & Sons, Inc.

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>





INEGI [Instituto Nacional de Estadística y Geografía] (2021). Censo de Población y Vivienda 2020, Bases de datos del Cuestionario Ampliado. Disponible en <https://www.inegi.org.mx/programas/ccpv/2020/default.html>

Machado, A. et al. (2010). Distribution of total height, transverse area and individual volume for *Araucaria angustifolia* (Bert.) O. Kuntze, *CERNE*, 16(1), 12-21.

Marquina B., Juárez R., O., Castells G., E. (2015). Indicadores sintéticos de bienestar social: una aplicación para los municipios del estado de Guerrero, México, *Revista Mexicana de Ciencias Agrícolas*, 1, 2015, 293-305. Disponible en <https://www.redalyc.org/articulo.oa?id=263139243039>

Mavrou, I. (2015). Análisis factorial exploratorio: Cuestiones conceptuales y metodológicas. *Revista Nebrija de Lingüística Aplicada*, 19, 71-80. Disponible en <https://doi.org/10.26378/rnlael019283>

Nayak, P. and Mishra, Sudhanshu K., Efficiency of Pena's P2 Distance in Construction of Human Development Indices (May 25, 2012). Disponible en SSRN: <https://ssrn.com/abstract=2066567> or <http://dx.doi.org/10.2139/ssrn.2066567>

Pena, T. B. (2009). La medición del Bienestar Social: una revisión crítica, *ESTUDIOS DE ECONOMÍA APLICADA*, 27(2), 299-324. Disponible en <https://www.redalyc.org/articulo.oa?id=30117056001>

Perez-Luque, A. J., Moreno, R., Perez-Perez, R. y Bonet, F. J. (2015). Package 'p2distance'. Disponible en <https://cran.r-project.org/web/packages/p2distance/p2distance.pdf>

Somarriba, N. y Pena, B. (2009). Synthetic Indicators of Quality of Life in Europe. *Social Indicators Research*. Recuperado de: doi 10.1007/s11205-008-9356-y

Somarriba, A. N., Zarzosa, E. P. y Pena, T. B. (2013). La calidad de vida en la Unión Europea. Un análisis temporal por medio de indicadores sintéticos. Documento presentado en la XXXIX Reunión de Estudios Regionales, Oviedo, España. Disponible en <https://old.reunionesdeestudiosregionales.org/Oviedo2013/htdocs/pdf/p851.pdf>

Wright, T. (2014). A Simple Method of Extract Optimal Allocation under Stratification with any Mixed Constraint Patterns. U.S. Census Bureau. Disponible en <https://www.census.gov/library/working-papers/2014/adrm/rrs2014-07.html>

Zarzosa, E. P. (1996). Aproximación a la medición del bienestar social. Idoneidad del indicador sintético "Distancia-P," (Aplicación al caso español). *Cuadernos de economía: Spanish Journal of Economics and Finance*, 24(68), 139-163. Recuperado de <https://repositorio.uam.es/handle/10486/5105>

\_\_\_\_\_. (2009). Estimación de la pobreza en las comunidades autónomas españolas, mediante la distancia DP2 de Pena. *Estudios de Economía Aplicada*, 27 (2), 397-415.

