

El Índice de Marginación de DP₂ no es un indicador cardinal en el tiempo y el espacio

Óscar Peláez Herreros¹

Resumen

El Índice de Marginación (IM) calculado por el Consejo Nacional de Población (Conapo) durante más de tres décadas se ha convertido en una herramienta muy utilizada en los análisis regionales ya que facilita información regular y sistematizada sobre las condiciones de vida de la población residente en los estados, municipios, localidades y áreas urbanas de México. A pesar de sus bondades, el IM recibió críticas por ofrecer medidas relativas, no directamente comparables en el tiempo. Para atender esta cuestión, el Conapo modificó la técnica de cálculo del IM para la versión de 2020. En concreto, dejó de utilizar el tradicional análisis de componentes principales para remplazarlo por el método DP₂ que, en palabras del Conapo, “permite valorar las brechas temporales y territoriales de manera directa”.

La comparabilidad de los indicadores sintéticos se relaciona con el postulado de homogeneidad, que ha de verificarse para que un indicador sintético pueda ser medible según escalas cardinales. El objetivo de este texto es demostrar que la técnica de DP₂ sólo genera resultados cardinales en el tiempo o en el espacio, pero no en ambas dimensiones a la vez, salvo que las desviaciones estándar y las relaciones entre variables sean constantes, algo que no ocurre. De manera más específica, el texto demuestra que el procedimiento seguido por el Conapo da lugar a un IM que verifica el postulado de homogeneidad y ofrece resultados cardinales para cada año de manera individual pero no para un conjunto de varios años. En consecuencia, se concluye que los resultados del IM calculado por DP₂ son comparables en el espacio, pero no a lo largo del tiempo. El conocimiento de esta cuestión, hasta ahora ignorada, puede ser relevante para ciertos análisis, como los que se proponen para este Encuentro de cambios en los escenarios territoriales a raíz de la reconfiguración del orden mundial.

Además, se comprueba que al sustituir la base de referencia del IM del Conapo por la de marginación nula, se recupera el sentido tradicional de los resultados del IM (mayor valor del índice indica mayor marginación) y se consigue la cardinalidad en los valores observados y no sólo en las distancias, lo que permite comparar regiones en términos de niveles de carencias. En el IM de 2020 del Conapo mayores valores indican menor marginación y las comparaciones se realizan como distancias al peor escenario hipotético del periodo 2010-2020. El cambio de base de referencia que se propone facilita la interpretación de los resultados e incrementa las posibilidades de análisis sin menoscabo de las propiedades del indicador.

Conceptos clave: Comparaciones interregionales, Postulado de homogeneidad, Marginación nula.

¹ Doctor, El Colegio de la Frontera Norte, opelaez@colef.mx

Introducción

La primera propuesta de un “índice de marginación” (IM) fue realizada en 1982 por la Coordinación General del Plan Nacional de Zonas Deprimidas y Grupos Marginados (Coplamar). El índice tomaba el nombre de la institución y del concepto definido en el documento de Coplamar (1977: 10) para “caracterizar aquellos grupos que han quedado al margen de los beneficios del desarrollo nacional y de los beneficios de la riqueza generada, pero no necesariamente al margen de la generación de la riqueza ni mucho menos de las condiciones que la hacen posible”. El índice en cuestión se calculaba mediante el método de componentes principales a partir de 19 variables observadas en las entidades federativas y municipios del país (Coplamar, 1982: 26-28).

A partir de la década de 1990, el proyecto fue retomado por el Consejo Nacional de Población (Conapo), que empezó a calcular el IM para los mismos niveles de desagregación territorial y con la misma técnica de agregación, pero utilizando sólo nueve variables e incorporando un proceso de estandarización y de estratificación formal de los resultados. Desde el año 2000, el IM también se estima para las localidades y las áreas geoestadísticas urbanas.

Con el transcurso del tiempo, este indicador sintético de las condiciones de vida de la población acabó convirtiéndose en una herramienta de apoyo para muchos análisis regionales, se le empezó a exigir más de lo que permitía su diseño y se criticó que sólo ofreciera medidas relativas no directamente comparables en el tiempo. Por ejemplo, Cárdenas (2010) explicaba que el IM calculado mediante análisis de componentes principales era una medida ordinal, que sólo ordenaba áreas geográficas, y no permitía hacer comparaciones cuantitativas en un momento del tiempo o a lo largo de éste; luego vinculaba la posibilidad de estas comparaciones con el hecho de que el índice pudiera expresarse en escala cardinal, en vez de ordinal. Con este propósito, para la versión de 2020, el Conapo ha reemplazado la técnica de cálculo del IM sustituyendo el tradicional análisis de componentes principales por el método de distancia de Pena (1977), denominado método de medición de distancia P_2 o método DP_2 , que “permite valorar las brechas temporales y territoriales de manera directa” (Conapo, 2021: 1).

El objetivo del presente texto es demostrar que la técnica de DP_2 sólo genera resultados cardinales en el tiempo o en el espacio, pero no en ambas dimensiones a la vez, salvo que las desviaciones estándar y las relaciones entre variables sean constantes, algo que no ocurre en el caso del IM. En concreto, se demuestra que el procedimiento utilizado por el Conapo genera resultados cardinales para las distancias en el espacio, pero no en el tiempo. Además, se comprueba que al cambiar la base de referencia utilizada por el Conapo se puede conseguir que el IM sea cardinal en las observaciones y no sólo en las distancias, lo que permite comparar regiones en términos de niveles de carencias, facilitando el análisis de los distintos escenarios territoriales.

Para esto, a continuación se describe el método de DP_2 . Seguidamente, se verifica si la propuesta del Conapo cumple el postulado de homogeneidad y en qué casos los resultados del IM se pueden interpretar en sentido cardinal. En el cuarto apartado se comprueba el efecto de un cambio de base de referencia. Por último, se presentan las conclusiones.

El método de los IM de 2020: DP₂

Conapo (2021) explica que los IM por estados y municipios (así como para localidades, áreas geoestadísticas básicas y colonias) correspondientes al año 2020 se estiman por el método DP₂. Destaca las múltiples ventajas de este procedimiento que “permite realizar comparaciones espaciales y temporales; admite variables expresadas en diferentes unidades de medida; evita la duplicación de información y tiene imparcialidad en el esquema de ponderación” (Conapo, 2021: 8). También menciona que el método verifica las propiedades matemáticas de existencia, determinación, monotonía, unicidad, invariancia, homogeneidad, transitividad, exhaustividad, aditividad, invariancia respecto a la base de referencia y neutralidad, como enumera Zarzosa (1996: 143).

El método fue propuesto originalmente por Pena (1977), que define el indicador sintético DP₂ como la suma de distancias ponderadas

$$DP_{2i} = \sum_{j=1}^n \frac{d_{ij}}{\sigma_j} (1 - R_{j,j-1,\dots,1}^2), \quad (1)$$

donde $d_{ij}=|x_{ij}-x_j|$ es la distancia entre el territorio i -ésimo y la base de referencia en la variable j -ésima, x_{ij} es la observación de la variable j -ésima en el territorio i -ésimo, x_j es la referencia para la variable j -ésima, σ_j es la desviación estándar de la variable j -ésima, y $R_{j,j-1,\dots,1}^2$ es el coeficiente de determinación de la regresión de X_j sobre $X_{j-1}, X_{j-2}, \dots, X_1$, siendo $R_1^2 = 0$. Al dividir las distancias por las desviaciones estándar, el procedimiento consigue que las distancias se midan en términos estandarizados. Con esta operación, las distancias quedan expresadas en unidades abstractas, lo que permite agregarlas aunque procedan de variables inicialmente medidas en unidades diferentes. Al mismo tiempo, σ_j puede interpretarse como una ponderación que penaliza a las variables con mayor dispersión (Pena, 1977: 92). Para el IM, esta segunda interpretación es más relevante que la primera ya que sus nueve variables componentes están expresadas en porcentaje, pero no todas tienen el mismo promedio ni la misma dispersión. Por su parte, $(1-R_{j,j-1,\dots,1}^2)$ es el “factor de corrección”, que deduce de cada indicador d_{ij} la información redundante ya contenida en indicadores precedentes (Pena, 1977: 114). Cabe mencionar que el orden en que se incluyen los indicadores parciales en la expresión de DP₂ es decisivo para el resultado final ya que los coeficientes de correlación parcial se modifican al alterar el orden de entrada de las variables. Por ello Pena (1977: 117) describe un procedimiento que garantiza el cumplimiento del postulado de unicidad.

De manera más específica, para el IM de entidades federativas y municipios, las variables, x_j , que permiten cuantificar las distancias d_j son los porcentajes de: población de 15 años o más analfabeta, 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, viviendas particulares con hacinamiento, población que vive en localidades menores a 5,000 habitantes, y población ocupada con ingresos de hasta dos salarios mínimos. La mayoría de estas variables se han utilizado desde el IM de 1990 (Conapo, 1994), pero algunas han cambiado. La modificación más importante ha tenido lugar con el porcentaje de población de 15 años o más sin educación básica, que ha

reemplazado al porcentaje de población de 15 años o más sin primaria completa. Este cambio aumenta necesariamente el porcentaje de población que padece la carencia. El Conapo también ha modificado ligeramente la definición del porcentaje de ocupantes en viviendas particulares sin drenaje ni excusado. En la versión de 2020, incluye a las viviendas con tubería que va a dar a una barranca o grieta, y tubería que va a dar a un río, lago o mar. En ocasiones anteriores, estas viviendas eran consideradas con drenaje.

La base de referencia contra la que se calculan las distancias de cada territorio fue definida por el Conapo como el peor escenario del periodo 2010-2020 en cada variable. Los valores concretos de este peor escenario hipotético de las entidades federativas y municipios se presentan en el Cuadro 1. El problema de esta base de referencia es que no corresponde a ninguna situación real que facilite una interpretación intuitiva de las distancias consecuentes. La base de referencia elegida por el Conapo es una combinación de los peores datos registrados en distintos años y lugares. Por ejemplo, las referencias estatales mezclan el porcentaje de ocupantes en viviendas particulares con piso de tierra de Guerrero en 2010 (19.61%), el porcentaje de población ocupada con ingresos de hasta dos salarios mínimos observado en Chiapas en 2020 (85.57%), el porcentaje de ocupantes en viviendas particulares sin energía eléctrica de Oaxaca en 2010 (4.93%), etc. En principio, esto no es un inconveniente para el método, que genera las mismas distancias entre territorios al cumplir el postulado de invariancia respecto a la base de referencia. No obstante, como se demostrará, esta elección dificulta la comprensión de los resultados y limita la interpretación cardinal a las distancias.

Cuadro 1. Bases de referencia a nivel estatal y municipal para el IM 2020, %

Componentes del IM	Base estatal	Base municipal
% de población de 15 años o más analfabeta	17.91	66.74
% de población de 15 años o más sin educación básica	60.73	94.79
% de ocupantes en viviendas particulares sin drenaje ni excusado	19.84	89.90
% de ocupantes en viviendas particulares sin energía eléctrica	4.93	69.45
% de ocupantes en viviendas particulares sin agua entubada	29.79	99.74
% de ocupantes en viviendas particulares con piso de tierra	19.61	79.71
% de viviendas particulares con hacinamiento	53.90	83.24
% de población en localidades menores a 5,000 habitantes	61.51	100
% de población ocupada con ingresos de hasta 2 salarios mínimos	85.57	100

Fuente: Elaboración propia con datos de Conapo (2021: 7).

El postulado de homogeneidad y la cardinalidad

Pena (1977) relaciona la cardinalidad de los resultados de la ecuación (1) con el postulado de homogeneidad, que enuncia del siguiente modo:

“Homogeneidad. «La función matemática que define el indicador sintético en función de los indicadores simples debe ser una función homogénea de grado uno.»

Este postulado se pide para que el indicador sintético pueda ser medible según escalas cardinales. Así, si se duplica el valor de cada indicador parcial, el indicador sintético también debe duplicar su valor.” (Pena, 1977: 52).

En términos más generales, Castro (2002: 261) define el índice sintético como $I=f(I_1, I_2, \dots, I_n)$, y el postulado de homogeneidad como la igualdad $f(c.I_1, c.I_2, \dots, c.I_n)=c.f(I_1, I_2, \dots, I_n)$, donde c es una constante. Schuschny y Soto (2009: 15) describen la homogeneidad de grado uno del mismo modo.

Para los indicadores sintéticos de distancia, como la ecuación (1), Pena (1977: 100) considera que los indicadores simples, I_j , son los cocientes d_{ij}/σ_j , y explica que, si se multiplica cada d_{ij}/σ_j por una misma constante, la DP_{2i} queda multiplicada por dicha constante. Con ello, se verifica el postulado de homogeneidad y los resultados del indicador de distancia DP₂ pueden interpretarse en términos cardinales.

Aquí hay dos detalles relevantes que se suelen pasar por alto. El primero es que la función DP₂ es homogénea de grado uno en las distancias, d_{ij} , pero no en los valores observados de las variables, x_{ij} . Si en una unidad territorial todas las distancias son la mitad que en otra, el índice DP_{2i} de la primera unidad es la mitad del de la segunda. Cabe recordar que se trata de una medida de distancia. Por ello, su cardinalidad es en términos de distancias. No se puede exigir otra cosa.

Por ejemplo, los datos de 2020 que obtiene el Conapo para las entidades federativas sitúan a Chiapas con un valor de 11.99 y a Nuevo León con 23.44 (Conapo, 2022). Estas cifras indican que la distancia de Nuevo León a la base de referencia, es casi el doble que la de Chiapas.

El problema con las carencias que componen el indicador es que se expresan en relación a la base de referencia elegida, por lo que sólo se puede hablar de ellas en escala ordinal: la base de referencia representa las mayores carencias, por ello, como Chiapas está más cerca de la base, Chiapas padece mayores carencias que Nuevo León. ¿Cuánto mayores son esas carencias? ¿El doble? No. En principio, a partir de la ecuación (1), nada se puede decir en escala cardinal del nivel de carencias. DP₂ no es cardinal en las carencias observadas, sino en las distancias de esas carencias respecto a la base de referencia. En el siguiente apartado se retoma este inconveniente.

El segundo detalle relevante que se suele obviar es que la ecuación (1) no alude a ninguna dimensión temporal. Incluye las observaciones de las nueve carencias para cada entidad federativa o municipio, sus desviaciones estándar, los coeficientes de determinación y las referencias, pero sin diferencias de tiempo. Todo está medido en el mismo instante o es atemporal.

Somarriva (2008: 127) menciona que “trabajaremos con el caso de una comparación interespatial, siendo totalmente válido todo lo que aquí se asuma en el caso de las comparaciones temporales.” Si en la ecuación (1) los subíndices i de las unidades territoriales se sustituyen por subíndices t de momentos de tiempo, efectivamente, todo lo dicho sigue siendo válido y el indicador DP_{2t} verifica el postulado de homogeneidad en las distancias. El inconveniente es que, de este modo, se prescinde de la dimensión espacial. El indicador sintético calcularía las distancias de una sola unidad territorial respecto a la base de referencia en distintos momentos de tiempo.

Para comparar varias unidades territoriales en varios momentos de tiempo, como pretende el Conapo, no se opera con la ecuación (1), sino con la variante

$$DP_{2ti} = \sum_{j=1}^n \frac{d_{tij}}{\sigma_{tj}} (1 - R_{t;j,j-1,\dots,1}^2), \quad (2)$$

que incluye la dimensión espacial a través de los subíndices i y la dimensión temporal con los subíndices t . En realidad, así lo hace el Conapo. Calcula la distancia DP_2 para las $i=32$ entidades federativas (o para los municipios) en los años $t=2010, 2015$ y 2020 .

El problema que surge al considerar variaciones en el espacio y en el tiempo de manera conjunta es que la ecuación (2) no es homogénea de grado uno y sus resultados no se pueden interpretar en sentido cardinal, ni siquiera para las distancias estandarizadas. Si se multiplica cada d_{tij}/σ_{tj} por una misma constante, la DP_{2ti} queda multiplicada por dicha constante sólo si los factores de corrección $(1-R_{t;j,j-1,\dots,1}^2)$ no cambian con el transcurso del tiempo. En el caso de que cambien, DP_{2ti} queda multiplicada por la constante y por el cambio en cuestión, de manera que se incumple la igualdad $f(c.I_1, c.I_2, \dots, c.I_n) = c.f(I_1, I_2, \dots, I_n)$.

Pena (1977: 118) recuerda que las propiedades del método de DP_2 se basan en un conjunto de condiciones entre las que se encuentra la invariancia de σ_j y $R_{j,j-1,\dots,1}^2$. También se refiere a las comparaciones en el tiempo e indica que

“la comparación directa sólo podría hacerse en el caso de que fuesen invariantes en el tiempo los parámetros σ y R^2 , supuesto, claro está, que se tome siempre la misma base de referencia y que ésta satisfaga las condiciones tantas veces enunciadas de ser igual o menor que el mínimo o igual o mayor que el máximo.” (Pena, 1977: 121).

El Conapo toma una misma base de referencia válida para calcular los IM por DP_2 de los años 2010, 2015 y 2020, pero no tiene en cuenta que los parámetros cambian en el tiempo. Como se observa en el Cuadro 2, el factor corrector de cada variable se modifica cada año, aunque también es cierto que, en esa inconstancia, hay casos de estabilidad. Algunos factores, como el correspondiente a las viviendas con piso de tierra, apenas varían durante la década analizada. Sin embargo, hay otros que modifican sustancialmente su valor, como el del porcentaje de población de 15 años o más sin educación básica, que para las entidades federativas se reduce en un tercio, o el de la población ocupada con ingresos de hasta dos salarios mínimos, que casi duplica su valor.

En estas condiciones, un estado o un municipio que entre 2010 y 2020 exactamente duplique su distancia estandarizada a la base de referencia en los nueve indicadores, no necesariamente duplica su DP_2 . El IM calculado como DP_2 no es una función homogénea de grado uno a lo largo del tiempo y sus resultados no pueden interpretarse a modo de escala cardinal.

Cuadro 2. Factores de corrección del IM por entidades federativas y municipios, 2010-2020

Componentes del IM	Entidades federativas			Municipios		
	2010	2015	2020	2010	2015	2020
% de población de 15 años o más analfabeta	1	1	1	1	1	1
% de población de 15 años o más sin educación básica	0.204	0.162	0.138	0.301	0.408	0.390
% de ocupantes en viviendas particulares sin drenaje ni excusado	0.540	0.595	0.573	0.752	0.757	0.638
% de ocupantes en viviendas particulares sin energía eléctrica	0.460	0.432	0.504	0.792	0.737	0.740
% de ocupantes en viviendas particulares sin agua entubada	0.143	0.207	0.159	0.709	0.742	0.791
% de ocupantes en viviendas particulares con piso de tierra	0.214	0.208	0.205	0.652	0.548	0.584
% de viviendas particulares con hacinamiento	0.256	0.306	0.280	0.557	0.578	0.504
% de población en localidades menores a 5,000 habitantes	0.242	0.276	0.272	0.525	0.601	0.568
% de población ocupada con ingresos de hasta 2 salarios mínimos	0.225	0.294	0.405	0.477	0.510	0.444

Fuente: Cálculos propios con datos de Conapo (2022).

La situación es aún más clara cuando se analiza la posible homogeneidad de las distancias, d_{tj} , en vez de las distancias estandarizadas, d_{tj}/σ_{tj} . Entonces, han de mantenerse constantes en el tiempo los factores de corrección $(1-R_{t,j,j-1,\dots,1}^2)$ y las desviaciones estándar σ_{tj} . Como se observa en el Cuadro 3, el conjunto de estos elementos está aún más lejos de poder considerarse invariante en el tiempo. Hay elementos que a nivel municipal casi son constantes todo el periodo, como ocurre con las ponderaciones de los porcentajes de viviendas particulares con hacinamiento y de población en localidades menores a 5,000 habitantes. Estos mismos casos también se mantienen relativamente estables entre las entidades federativas, pero con ritmos de crecimiento ligeramente mayores (Cuadro 4). En contraste, la ponderación que el procedimiento asigna a otras distancias más que se duplica en apenas 10 años. Así ocurre entre las entidades federativas con los porcentajes de ocupantes en viviendas particulares sin drenaje ni excusado y de población ocupada con ingresos de hasta dos salarios mínimos, y en los municipios y entidades federativas con los porcentajes de ocupantes en viviendas particulares sin energía eléctrica y sin agua entubada (Cuadro 4). Con tasas de variación superiores a 100% en una década, es difícil aceptar la invariancia de σ y R^2 en el tiempo.

Cuadro 3. Ponderaciones $(1-R_{t;j,j-1,\dots,1}^2)/\sigma_{tj}$ del IM por entidades federativas y municipios, 2010-2020

Componentes del IM	Entidades federativas			Municipios		
	2010	2015	2020	2010	2015	2020
% de población de 15 años o más analfabeta	0.239	0.285	0.320	0.103	0.116	0.131
% de población de 15 años o más sin educación básica	0.025	0.021	0.019	0.021	0.028	0.028
% de ocupantes en viviendas particulares sin drenaje ni excusado	0.140	0.220	0.313	0.080	0.105	0.121
% de ocupantes en viviendas particulares sin energía eléctrica	0.347	0.529	0.859	0.127	0.205	0.267
% de ocupantes en viviendas particulares sin agua entubada	0.020	0.050	0.052	0.040	0.064	0.086
% de ocupantes en viviendas particulares con piso de tierra	0.046	0.059	0.057	0.055	0.062	0.065
% de viviendas particulares con hacinamiento	0.038	0.050	0.046	0.045	0.051	0.048
% de población en localidades menores a 5,000 habitantes	0.015	0.017	0.017	0.015	0.017	0.016
% de población ocupada con ingresos de hasta 2 salarios mínimos	0.019	0.028	0.043	0.024	0.030	0.037

Fuente: Cálculos propios con datos de Conapo (2022).

Cuadro 4. Variación porcentual de las ponderaciones de DP₂ en el periodo 2010-2020, %

Componentes del IM	$1-R_{t;j,j-1,\dots,1}^2$		$(1-R_{t;j,j-1,\dots,1}^2)/\sigma_{tj}$	
	Entidades federativas	Municipios	Entidades federativas	Municipios
% de población de 15 años o más analfabeta	0%	0%	34%	27%
% de población de 15 años o más sin educación básica	-33%	29%	-25%	35%
% de ocupantes en viviendas particulares sin drenaje ni excusado	6%	-15%	123%	50%
% de ocupantes en viviendas particulares sin energía eléctrica	10%	-6%	147%	110%
% de ocupantes en viviendas particulares sin agua entubada	11%	12%	160%	115%
% de ocupantes en viviendas particulares con piso de tierra	-4%	-10%	25%	18%
% de viviendas particulares con hacinamiento	9%	-10%	21%	7%

% de población en localidades menores a 5,000 habitantes	12%	8%	14%	6%
% de población ocupada con ingresos de hasta 2 salarios mínimos	80%	-7%	122%	54%

Fuente: Cálculos propios con datos de Conapo (2022).

A diferencia de lo que ocurre con las comparaciones interanuales, el postulado de homogeneidad sí se verifica para las comparaciones entre territorios en un mismo año, ya que entonces los factores de corrección y las desviaciones estándar de las variables son los mismos: los correspondientes a ese año para todos los territorios que se estén comparando. Si un municipio presenta exactamente el doble de distancia respecto a la base de referencia en los nueve indicadores en comparación con otro municipio el mismo año, su DP₂ es exactamente el doble. La imprecisión surge al desplegar la dimensión temporal.

Pena (1977: 121) explica que no se trata de un problema matemático, sino de querer medir y establecer comparaciones en una realidad cambiante:

“el desarrollo o el nivel de vida es algo variable, tanto en su concepción como en su contexto. Los juicios de valor que subyacen en el fondo de ambos conceptos no son inmutables, sino que varían con el tiempo, lo que significa la imposibilidad de comparaciones exactas cuando las épocas son muy diferentes.” (Pena, 1977: 121).

Y concluye que,

“aunque no se logren resultados exactos, sí se pueden lograr buenas aproximaciones con tal de que los períodos a comparar no sean muy distantes y de que se tome siempre la misma base de referencia.” (Pena, 1977: 121).

Esto es lo que obtiene el Conapo: un indicador comparable en el espacio, pero sólo aproximadamente comparable en el tiempo, siendo esa aproximación poco precisa porque las desviaciones estándar y las correlaciones cambian bastante en diez años.

No es correcto lo que afirman Somarriba et al. (2013: 11) cuando aplican la técnica de DP₂ al análisis de la calidad de vida en varios países europeos en 2007 y 2011. En concreto, argumentan que:

“Debido a las propiedades del indicador sintético DP₂, es posible interpretar cardinalmente las distancias entre cada dos países. [...]

Además, dado que se ha calculado el indicador sintético utilizando la misma base de referencia para los dos períodos, también es posible interpretar cardinalmente las distancias entre los dos momentos temporales, para cada país.” (Somarriba et al., 2013: 11).

Esta última frase es errónea. Lo correcto es lo que explica Pena (1977: 121) y se ha demostrado en páginas anteriores. Para la aplicación que realizan Somarriba et al. (2013), que es análoga a la de Conapo (2022), es posible interpretar cardinalmente las distancias

entre dos territorios en un mismo año, pero no las distancias en el tiempo, que carecen de interpretación cardinal aunque se utilice la misma base de referencia. El cambio que tiene lugar en los factores de corrección impide que los resultados de la agregación se expresen en escala cardinal.

Como indica Pena (1977: 179), “las comparaciones en el tiempo hay que tomarlas con ciertas precauciones”. No sólo ha de permanecer constante la base de referencia, sino también, al menos, los parámetros R^2 . En la medida en que esto último no ocurre, las comparaciones inter-temporales sólo tienen un carácter aproximado.

Hay una forma de conservar la cardinalidad en el espacio y en el tiempo de manera simultánea al aplicar la técnica de DP₂. Consiste en establecer de inicio el supuesto de que las relaciones entre las variables y sus desviaciones estándar no cambian con el transcurso del tiempo (R^2 y σ constantes) y estimar esas relaciones y desviaciones con toda la información disponible sin importar el año de referencia. Así operan Somarriba (2008), Somarriba y Pena (2010) y Merino et al. (2012). Por ejemplo, Merino et al. (2012) explican que:

“Se propone calcular un único indicador sintético en el que intervienen todas las unidades territoriales, en los tres instantes de tiempo, como si fueran un único conjunto de observaciones, de forma que cada unidad territorial se considera diferente para cada período.” (Merino et al., 2012: 270).

El inconveniente de este procedimiento es que las observaciones suelen demostrar que las relaciones entre variables cambian con el paso de los años. La agregación de las distancias con desviaciones estándar y factores de corrección fijos estaría introduciendo distorsiones que podrían evitarse calculando las variaciones y relaciones para cada año por separado, sólo que entonces los resultados no permiten comparaciones cardinales inter-temporales sino únicamente inter-territoriales, como se ha expuesto.

Pena (1977), Zarzosa (2012), Somarriba et al. (2013) y Conapo (2022) proceden de esta última forma. Eligen una base de referencia que mantienen constante, pero calculan las DP₂ para cada año por separado, con distintos factores de corrección y desviaciones estándar. Sus resultados cumplen el postulado de homogeneidad para las comparaciones inter-territoriales en el mismo momento de tiempo, pero carecen de una interpretación inter-temporal cardinal.

El efecto de un cambio de base de referencia: la base 0

El método de DP₂ verifica el postulado de invariancia respecto a la base de referencia, siempre que la base tenga un valor mayor para cada componente que el máximo observado, o menor, también para cada componente, que el mínimo (Pena, 1977: 62 y 118). Zarzosa (2009: 403) presenta condiciones más específicas para la validez de una base de referencia.

El cumplimiento de esta propiedad de invariancia permite que se pueda elegir cualquier base de referencia válida. Con ello, cambian las distancias respecto a la base, pero no las distancias relativas entre territorios. En consecuencia, lo más conveniente es elegir una base de referencia que ayude a interpretar los resultados obtenidos.

Conapo (2022) elige las bases de referencia del Cuadro 1, que combinan las peores condiciones de diversos territorios y años. Las bases son válidas según las condiciones de Zarzosa (2009: 403). El problema es que generan valores de, por ejemplo, 11.99 para Chiapas, 17.51 para Yucatán o 23.44 para Nuevo León en 2020 (Conapo, 2022). ¿Qué significan esos valores? Son distancias estandarizadas de cada unidad territorial a la base de referencia. Como ya se ha mencionado, dada su posible interpretación cardinal en términos de distancias, muestran que en 2020 Nuevo León se encontraba casi el doble de lejos de la base de referencia que Chiapas.

El punto que se argumenta es que hay otra base válida que aporta más información. En la propuesta del IM por DP₂, el Conapo se limitó a seguir a Zarzosa (2009), Zarzosa y Somarriba (2013) o a Somarriba et al. (2013) sin realizar una adaptación oportuna de la técnica al caso. Las variables que componen el IM tienen unas propiedades específicas que facilitan otra solución.

Pena (1977: 58) no define la base de referencia para DP₂ como una combinación de peores situaciones observadas, sino como un objetivo a alcanzar. En concreto, explica que, si la base de referencia “se presenta como el objetivo deseable tendríamos entonces un conocimiento de la «distancia» que nos falta por recorrer para llegar a la meta asignada.” Como las nueve variables del IM son porcentajes necesariamente comprendidos entre 0 y 100% que representan carencias, una meta deseable es que todas las variables se sitúen en 0%, esto es, que nadie padezca ninguna carencia, lo que Peláez (2017: 123) denomina “marginación nula”, que “se da cuando no se observa ninguna de las formas de exclusión del IM en ningún residente ni vivienda de un territorio”.

La ventaja de esta base de referencia es la siguiente. Dado que $x_{*j}=0 \forall j$, la expresión de las distancias queda reducida a $d_{ij}=|x_{ij}|$. Además, como los valores de las nueve variables del IM siempre son positivos, se puede prescindir del valor absoluto y escribir: $d_{ij}=x_{ij}$. La referencia de 0% $\forall j$ iguala las distancias a las carencias. Por ello, la ecuación (1) puede expresarse como

$$DP_{2b0i} = \sum_{j=1}^n \frac{x_{ij}}{\sigma_j} (1 - R_{j,j-1,\dots,1}^2) \quad (3)$$

Anteriormente se concluyó que la ecuación (1) era homogénea de grado uno en las distancias. De manera análoga, ahora se puede concluir que la ecuación (3) cumple el postulado de homogeneidad en los valores observados. Los resultados de la ecuación (3) se expresan en escala cardinal para las distancias y para las carencias, que en ella son lo mismo. Un territorio que presente exactamente el doble de valor en las nueve carencias en comparación con otro territorio el mismo año, tiene una DP_{2b0} que exactamente duplica a la del segundo territorio.

El IM para la base de marginación nula se puede calcular con los datos que utiliza el Conapo para su IM, siguiendo todo el procedimiento y sólo cambiando la base de referencia. No obstante, también se puede obtener directamente a partir del resultado del IM del Conapo. Para ello hay que tener en cuenta que, al sustituir la expresión de las distancias en la ecuación (1), ésta se reescribe como

$$DP_{2i} = \sum_{j=1}^n \frac{|x_{ij} - x_{*j}|}{\sigma_j} (1 - R_{j,j-1,\dots,1}^2) \quad (4)$$

Además, dado que la base de referencia tiene que cumplir las condiciones de Zarzosa (2009: 403), esto es, que $x_{ij} > x_{*j} \forall i, j$ o $x_{ij} < x_{*j} \forall i, j$, la ecuación (4) equivale a

$$DP_{2i} = \left| \sum_{j=1}^n \frac{x_{ij} - x_{*j}}{\sigma_j} (1 - R_{j,j-1,\dots,1}^2) \right| \quad (5)$$

En esta ecuación (5), el sumatorio se puede desagregar en dos, quedando

$$DP_{2i} = \left| \sum_{j=1}^n \frac{x_{ij}}{\sigma_j} (1 - R_{j,j-1,\dots,1}^2) - \sum_{j=1}^n \frac{x_{*j}}{\sigma_j} (1 - R_{j,j-1,\dots,1}^2) \right| \quad (6)$$

El primer sumatorio de esta ecuación (6) es el mismo de la ecuación (3); es, por tanto, el valor de DP₂ con base nula, DP_{2b0}. El segundo sumatorio no depende de *i*, tomando el mismo valor para todas las unidades territoriales *i* que se estén analizando. Puede representarse como una constante, *c*. El valor absoluto indica que el conjunto de los dos sumatorios debe ser positivo. Además, un sumatorio debe ser positivo y el otro negativo, y se sabe que DP_{2b0} mide las distancias hasta la mejor situación posible, esto es, a la inversa que la DP₂ que calcula el Conapo, que son distancias hasta un peor escenario hipotético. Teniendo esto en cuenta y reordenando términos, la ecuación (6) puede reescribirse como

$$DP_{2i} = c - DP_{2b0i} \quad (7)$$

La ecuación (7) es consecuencia de la verificación del postulado de invariancia respecto a la base de referencia, que cumple la técnica de DP₂. Para obtener el IM con base nula basta con cambiar de signo los valores del IM publicado por el Conapo y sumarles una constante que depende de las referencias del Conapo (Cuadro 1), de las desviaciones estándar y de los factores de corrección (Cuadro 3). Por estos valores, la constante *c* es distinta para cada conjunto de datos. Cambia según se trate de entidades federativas o de municipios y también cambia con el año que se analice. El Cuadro 5 contiene estas constantes.

Cuadro 5. Valores de la constante *c* para cambiar la base de referencia del Conapo por la base de marginación nula, según año y nivel de desagregación territorial

	2010	2015	2020
Entidades federativas	16.4175	22.1491	27.1819
Municipios	40.9563	54.4459	63.8626

Fuente: Cálculos propios con datos de Conapo (2022).

El aumento de *c* con el transcurso del tiempo es el resultado de los cambios en σ y R^2 , porque la base de referencia es la misma para todo el periodo. Operando según la ecuación (7) con las constantes de las entidades federativas y el IM del Conapo, se obtienen los valores

del IM con base marginación nula que figuran en el Cuadro 6. Se puede proceder de forma análoga para los casi 2,500 municipios del país y obtener el listado correspondiente.

Cuadro 6. Índice de Marginación DP₂ con base marginación nula para las entidades federativas, 2010-2020

Entidades federativas	2010	2015	2020
Aguascalientes	4.3415	4.2651	4.9762
Baja California	3.8115	3.6849	5.8016
Baja California Sur	5.0674	4.8285	5.7086
Campeche	8.3748	8.4645	9.3769
Coahuila de Zaragoza	3.8114	3.7481	4.6363
Colima	4.7985	4.7166	5.6496
Chiapas	13.2867	13.6993	15.1833
Chihuahua	5.8868	5.3274	7.1667
Ciudad de México	2.8263	2.7509	4.0388
Durango	7.2635	7.2068	8.7092
Guanajuato	7.6141	7.0414	7.7625
Guerrero	14.9035	15.2207	16.1928
Hidalgo	9.0638	8.5903	9.1288
Jalisco	4.7408	4.5645	5.3668
Estado de México	5.3774	5.2338	6.3783
Michoacán de Ocampo	8.4469	8.2875	8.9011
Morelos	5.9895	6.0583	7.3680
Nayarit	7.5759	8.1187	9.6660
Nuevo León	3.1283	2.9229	3.7376
Oaxaca	12.8328	12.8426	13.9655
Puebla	8.7515	8.7090	9.4603
Querétaro de Arteaga	6.7804	5.8376	6.3438
Quintana Roo	5.8636	5.9180	6.5531
San Luis Potosí	8.4573	8.3635	8.4939
Sinaloa	6.0280	6.0838	6.6721
Sonora	4.8355	4.7421	5.7763
Tabasco	7.4797	7.1935	8.8495
Tamaulipas	4.8512	4.9583	6.1854
Tlaxcala	6.3817	6.2289	7.3112
Veracruz de Ignacio de la Llave	9.5742	9.7408	10.7678
Yucatán	8.9869	9.5399	9.6697
Zacatecas	7.1672	6.9296	7.6847

Fuente: Cálculos propios con datos de Conapo (2022).

El IM con base nula presenta algunas ventajas sobre el IM del Conapo basado en las peores condiciones de diversos años y territorios. En primer lugar, asigna mayores valores numéricos a las unidades territoriales con peores condiciones de vida, algo que cabe esperar de un índice de “marginación”. La versión del Conapo hace lo contrario, como si se tratara de un indicador de bienestar en vez de carencias.

El cambio de base no afecta a las distancias entre territorios. Las mismas 11.446 unidades que separan a Nuevo León de Chiapas en el IM del Conapo para 2020, separan a estas entidades en el índice de base nula. Las demás distancias entre estados o municipios tampoco registran cambios en ningún año. A consecuencia de ello, la estratificación de los resultados en grados de marginación tampoco se modifica.

Los valores del Cuadro 6 representan distancias, pero al estar referidas al cero también representan niveles de carencias. Por la propiedad de homogeneidad, que se verifica para las comparaciones inter-territoriales en el mismo momento de tiempo, puede afirmarse que Guerrero, con un IM DP_{2b0} de 16.1928 en 2020, se encuentra cuatro veces más lejos de la base de referencia que Nuevo León, que tiene un IM DP_{2b0} de 3.7376 ese mismo año (hasta aquí la interpretación es semejante a la del IM del Conapo, pero lo que sigue ya no), de manera que Guerrero padece cuatro veces más carencias que Nuevo León en 2020. Con la base nula, la técnica de DP_2 aplicada al IM devuelve resultados cardinales en las observaciones. Con cualquier otra base, la cardinalidad sólo es en términos de distancias. Pero ocurre que con la base de marginación nula las distancias equivalen a las observaciones y se pueden interpretar en el sentido señalado.

La última ventaja de la base de marginación nula es que ayuda a entender que las comparaciones inter-temporales de DP_2 carecen de fundamento en el periodo analizado. Como se puede comprobar en el Cuadro 6, las cifras de DP_2 no tienden a aproximarse a 0 con el transcurso del tiempo, sino que incluso tienden a alejarse ligeramente (por ejemplo, Guanajuato: 7.61 en 2010, 7.04 en 2015 y 7.76 en 2020). Sin embargo, las estimaciones de Conapo (2022) se alejan claramente de su base de referencia (el mismo ejemplo de Guanajuato: 8.80 en 2010, 15.11 en 2015 y 19.42 en 2020). ¿A qué se deben estas diferencias si lo único que distingue ambos casos es el cambio de base y DP_2 verifica el postulado de invariancia respecto a la base de referencia? Las diferencias se deben a la violación de la invariancia de σ y R^2 .

Como se observa en los Cuadros 3 y 4, el conjunto $(1 - R_{t,j,j-1,\dots,1}^2) / \sigma_{tj}$ aumenta sustancialmente con el transcurso de los años. Esto invalida la propiedad de homogeneidad en el tiempo. Esas secuencias de números carecen de cualquier posible interpretación cardinal. Sólo tienen sentido en el contexto en que se generan:

- Cuando la base de referencia representa la peor situación, los territorios se alejan de ella, las distancias d_{tj} aumentan y además se ven multiplicadas por ponderaciones $(1 - R_{t,j,j-1,\dots,1}^2) / \sigma_{tj}$ en aumento. El resultado indica grandes avances en el distanciamiento respecto a la peor situación posible.
- En cambio, cuando la base de referencia representa la mejor situación, los territorios se acercan a ella, las distancias d_{tj} disminuyen, pero se ven compensadas por ponderaciones $(1 - R_{t,j,j-1,\dots,1}^2) / \sigma_{tj}$ en aumento. El resultado es de mínimo avance o incluso retroceso en el acercamiento a la mejor situación posible.

Ninguna de estas interpretaciones es correcta en términos absolutos, ni siquiera de forma aproximada, porque el cambio en σ y R^2 es sustancial, hasta el punto de dar lugar a resultados contrarios dependiendo de la base de referencia que se elija.

Conclusiones

Se ha demostrado que el IM que el Conapo calcula por DP₂ cumple el postulado de homogeneidad en el espacio, pero no en el tiempo. Es, por ello, una medida cardinal en las distancias para las comparaciones inter-territoriales, pero no permite el análisis de la evolución de esas distancias en el tiempo ni siquiera de forma aproximada. Los resultados cambian sustancialmente según la base de referencia que se elija. Esto es debido a que el cálculo del IM de cada año utiliza las desviaciones estándar y las correlaciones de ese año, que difieren de las de otros años, de manera que las comparaciones inter-temporales violan el supuesto de invariancia de los parámetros σ y R^2 que la técnica de DP₂ requiere para el cumplimiento de sus propiedades.

Además, se ha demostrado que el uso de la base de referencia de marginación nula, en sustitución de la elegida por el Conapo, permite recuperar el sentido tradicional de los valores del IM (mayor valor numérico vuelve a corresponder con mayor marginación) e interpretar las observaciones (las carencias) en sentido cardinal, algo que el IM del Conapo no consigue ya que su cardinalidad se limita a las distancias respecto a su base. Esta posibilidad de comparar directamente los niveles de carencias abre nuevas vías de análisis. Finalmente, el cambio de base de referencia también ha servido para evidenciar las grandes diferencias que puede haber al interpretar los valores de DP₂ a lo largo del tiempo.

Por todo lo mencionado, cabe concluir que la pretensión del Conapo de realizar comparaciones inter-temporales con su nuevo IM carece de sustento y debe descartarse. Por otra parte, cuando la técnica de DP₂ se adapta a las particularidades del IM como se ha detallado en las páginas anteriores, surgen posibilidades de análisis de las diferencias interregionales que en la actualidad se están desaprovechando. Ambos aspectos son relevantes para comprender las limitaciones y alternativas que ofrecen las cifras del nuevo IM de DP₂ como instrumento para la descripción, caracterización y comparación de los escenarios territoriales que se analizan en este Encuentro a propósito de la reconfiguración del orden mundial.

Referencias

- Cárdenas, Ó. J.**, (2010) “Cardenalización del índice de marginación: una metodología para evaluar la eficiencia del gasto ejercido en el Ramo 33” en *EconoQuantum*, vol. 7, núm. 1, pp. 41-66.
- Castro, J. M.**, (2002) *Indicadores de desarrollo sostenible urbano: una aplicación para Andalucía*. Tesis de doctorado. España, Universidad de Málaga.
- Conapo**, (1994) *Desigualdad regional y marginación municipal en México, 1990*. México, Consejo Nacional de Población / Comisión Nacional del Agua.
- Conapo**, (2021) “Índice de marginación por entidad federativa y municipio 2020: nota técnico-metodológica” en *Consejo Nacional de Población*, Ciudad de México, disponible en:
https://www.gob.mx/cms/uploads/attachment/file/685354/Nota_tecnica_IMEyM_2020.pdf [22 de julio de 2022].

- Conapo**, (2022) “Índices de marginación 2020” en *Consejo Nacional de Población*, Ciudad de México, disponible en: <https://www.gob.mx/conapo/documentos/indices-de-marginacion-2020-284372> [7 de julio de 2022].
- Coplamar**, (1977) *Bases para la acción. 1977-1982*. México, Coordinación General del Plan Nacional de Zonas Deprimidas y Grupos Marginados.
- Coplamar**, (1982) *Necesidades esenciales en México: situación actual y perspectivas al año 2000, vol. 5: Geografía de la marginación*. México, Siglo XXI.
- Merino, M. C.; Somarriba, N. y A. M. Negro**, (2012) “Un análisis dinámico de la calidad del trabajo en España. Los efectos de la crisis económica” en *Estudios de Economía Aplicada*, vol. 30, núm. 1, pp. 261-282.
- Peláez, Ó.**, (2017) “La marginación a lo largo del tiempo: cálculo del Índice de Marginación Absoluta (IMA) para las entidades federativas de México, 1970-2010” en *Economía: Teoría y Práctica*, núm. 46, pp. 115-137.
- Pena, B.**, (1977) *Problemas de la medición del bienestar y conceptos afines. (Una aplicación al caso español)*. España, Instituto Nacional de Estadística.
- Schuschny, A. y H. Soto**, (2009) *Guía metodológica: Diseño de indicadores compuestos de desarrollo sostenible*. Chile, Comisión Económica para América Latina y el Caribe.
- Somarriba, N.**, (2008) *Aproximación a la medición de la calidad de vida social e individual en la Europa Comunitaria*. Tesis de doctorado. España, Universidad de Valladolid.
- Somarriba, N. y B. Pena**, (2010) “Un análisis dinámico de la calidad de vida y de la convergencia en Europa” en *Anales de Estudios Económicos y Empresariales*, vol. 20, pp. 283-324.
- Somarriba, N.; Zarzosa, P. y B. Pena**, (2013) “La calidad de vida en la Unión Europea. Un análisis temporal por medio de indicadores sintéticos” en *XXXIX Reunión de Estudios Regionales*, 21 y 22 de noviembre de 2013, Oviedo.
- Zarzosa, P.**, (1996) “Aproximación a la medición del bienestar social. Idoneidad del indicador sintético ‘Distancia-P₂’. (Aplicación al caso español)” en *Cuadernos de Economía*, vol. 24, pp. 139-163.
- Zarzosa, P.**, (2009) “Estimación de la pobreza en las comunidades autónomas españolas, mediante la distancia DP₂ de Pena” en *Estudios de Economía Aplicada*, vol. 27, núm. 2, pp. 397-415.
- Zarzosa, P.**, (2012) “The social welfare in Spain before the crisis: territorial and chronological analysis” en *International Journal of Advances in Management and Economics*, vol. 1, núm. 4, pp. 165-171.
- Zarzosa, P. y N. Somarriba**, (2013) “An assessment of social welfare in Spain: territorial analysis using a synthetic welfare indicator” en *Social Indicators Research*, vol. 111, núm. 1, pp. 1-23.