

Debate*

Confiabilidad de la confiabilidad

Arturo Blancas Espejo**

En el número anterior de esta revista (Vol. 6, núm. 1, enero-abril, 1991) apareció el artículo "Confiabilidad de los resultados preliminares del XI Censo General de Población y Vivienda de 1990" elaborado por Rodolfo Corona Vázquez.

Como Director de Censos Nacionales, del Instituto Nacional de Estadística, Geografía e Informática (INEGI), tuve oportunidad de participar directamente en la realización del último censo de México, por lo que el título del mencionado artículo me resultó de gran interés y lo leí con la esperanza de encontrar elementos que permitieran dimensionar el magnífico esfuerzo que llevó a cabo el país durante el censo de 1990.

Mis expectativas no fueron cubiertas, ya que el mencionado artículo incurre, desde mi punto de vista, en errores de método que empobrecen los resultados que se presentan, al grado que el porcentaje de omisión censal que ahí se calcula, mide una proporción diferente a la que intenta estimar.

Por lo anterior, y buscando orientar a los usuarios de la información censal, para que ubiquen correctamente la calidad de los diversos productos que se han generado sobre el XI Censo General de Población y Vivienda de 1990, me resulta imperioso hacer algunas reflexiones y comentarios sobre el mencionado artículo.

Para facilitar la presentación de los comentarios seguiremos la estructura que Rodolfo Corona estableció en su documento y haremos comentarios para cada una de las secciones que lo integran. Al final, y bajo el nombre de Conclusiones, se presenta un resumen de los puntos más importantes aquí comentados.

Utilidad y exactitud de un censo de población

Sería conveniente agregar en el primer párrafo de esta sección que la mencionada fotografía estadística cubre a todos y cada uno

* Los textos de esta sección se publican en su versión original; no pasan por revisión de estilo.

** Director de Censos Nacionales del INEGI.

de los habitantes de un país, de lo contrario se tendría una equivalencia entre un censo y una encuesta de cobertura nacional.

Los Censos de Población y Vivienda que se han realizado en nuestro país, consideran como unidad de observación o de análisis, a la vivienda, al hogar o al individuo. Para cada una de éstas se registran características (variables), las cuales se miden en diversas unidades, como pueden ser años, pesos, etc. En ningún caso se ha utilizado al individuo como unidad de medida; el individuo no es equivalente a un centímetro o a un kilogramo como aparentemente se establece en el texto.

Realizar un censo implica una inversión considerable de recursos y de tiempo, de ahí que los países que los llevan a cabo, siempre los consideran la fuente de información más importante del Sistema Nacional de Información.

Una de las principales aportaciones de los censos de población a los sistemas de información de los países, es que proporcionan datos que pueden utilizarse para la construcción de marcos muestrales de viviendas, a partir de los cuales se pueden realizar encuestas probabilísticas, cuyas estimaciones tienen confiabilidad y precisión medibles. Este aspecto, que es de gran relevancia, no es equivalente a decir, como se hace en el texto, que el censo es importante porque proporciona una base cuantitativa útil "para producir otras estadísticas (como las derivadas de encuestas por muestreo)".

En el texto se menciona que la importancia de un censo está supeditada a su veracidad, lo cual no es privativo de los proyectos censales, sino que también es válido para los demás procedimientos de generación de información estadística.

Las condiciones que se presentan, para que un censo genere cifras exactas; son las mismas que deben respetarse para el caso de las encuestas y de los registros administrativos, por lo que no constituyen aspectos exclusivos de los proyectos censales.

Habría que agregar, a lo presentado por el autor, que los censos de población, y en particular el XI Censo General de Población y Vivienda de 1990, nos permiten saber cuántos somos, cómo somos y una característica básica del censo es el poder saber dónde estamos. Los resultados del Censo de 1990, proporcionan información a nivel nacional, por entidad federativa, por municipio, por localidad, y para las localidades de más de 2 500 habitantes por grupo de manzanas. Esta propiedad de ubicar los montos poblacionales y sus características, a niveles geográficos reducidos, es lo que hace del Censo de Población y Vivienda, al menos en nuestro país, la columna vertebral del Sistema Nacional de Información Estadística.

Evaluación de un censo de población

Seguramente el autor estará de acuerdo en que para detectar las causas de un error, primero debemos conocer la existencia del mismo. Consecuentemente, sugerimos un cambio de orden en las etapas que propone para llevar a cabo la evaluación censal. Necesariamente la "Evaluación Numérica" debe realizarse antes de la denominada "Evaluación del Proceso de Generación".

El procedimiento propuesto para la "Evaluación Numérica" compara los resultados censales con los de otras fuentes, a las cuales se les confiere, sin demostrarlo, una calidad superior a la de las cifras censales; es decir, de antemano se descalifica la calidad de éstas. Consecuentemente el procedimiento que se propone no establece evaluaciones de carácter absoluto sino relativo; es decir, todo depende de la calidad que realmente tengan las cifras que sirven de punto de comparación, sobre la cual, como ya se mencionó, no se aporta evidencia alguna. Si se aplica el método propuesto, pero ahora considerando como de mejor calidad a las cifras del censo, se tendría que los resultados que aparecen, en el artículo que comentamos, como negativos para el censo lo serían entonces para las otras fuentes.

Lo expresado en el párrafo anterior resulta fundamental. El método propuesto genera resultados de carácter relativo y no se tiene la certeza de que los elementos que proporciona dimensionen correctamente la cobertura censal.

Las evaluaciones de los censos mexicanos

En esta sección se afirma que mientras más pequeña es la zona geográfica de que se trate, más aumenta la complejidad del fenómeno demográfico correspondiente. Si esto fuera así, sería más fácil entender demográficamente hablando, el comportamiento del total de los habitantes del estado de México, que el comportamiento de los residentes de una manzana en alguna colonia del Distrito Federal.

El fenómeno demográfico, en cualquier espacio geográfico, queda determinado por las entradas (nacimientos e inmigraciones), y las salidas (defunciones y emigraciones) que llevan a cabo los integrantes del grupo poblacional que se estudia. En cualquier espacio geográfico se hace patente la validez de la ecuación compensadora, tan usada en el método demográfico.

Lo que sí resulta complejo en pequeñas zonas geográficas, es por una parte disponer de información estadística de calidad, en

especial en lo referente a la migración, y por la otra que los métodos y procedimientos utilizados en demografía, son vigorosos cuando se aplican a grandes agregados de población, y presentan debilidades al utilizarse para poblaciones de montos reducidos.

En el texto se afirma que la falta de evaluaciones censales a nivel regional ha originado el supuesto de que el grado de cobertura global de los censos es válido para todas las regiones de la república. Al respecto se debe comentar que el establecimiento de un supuesto queda totalmente en manos del investigador de que se trate, quien deberá evaluar las ventajas y desventajas de considerar válida determinada situación. En ningún caso el investigador se ve obligado a establecer un supuesto, que en este caso, como el mismo texto indica, no se cumple.

En el texto se afirma que "el punto de partida de las proyecciones demográficas debe ser una serie de cifras que de la mejor manera posible represente la real población por sexo y edad; por ello se usan como base los datos censales, pero corregidos o ajustados".

Considero que sería de gran utilidad, para los sistemas de información estadística de todos los países, que el autor expusiera en forma explícita, cuál es esa "mejor manera posible" a que hace referencia, mediante la cual se conoce el valor real de la población; aparentemente, con la aplicación de ese procedimiento ya no sería necesario realizar censos de población, lo cual implica un enorme ahorro de tiempo y recursos.

Se afirma que las evaluaciones censales que se hacen actualmente son mucho más precisas que las realizadas en el pasado, en especial por la aplicación de un "procedimiento llamado conciliación censal". Sobre este punto se debe apuntar que la conciliación censal, que por cierto la mayoría de los especialistas llama conciliación demográfica, es un procedimiento que genera escenarios donde las componentes de la dinámica demográfica, estimadas a través de diferentes mecanismos, y las cifras censales corregidas, es decir, también estimadas, muestran comportamientos consistentes. El método sería de gran utilidad si permitiera obtener una sola solución que estableciera la consistencia entre los diversos elementos que se hacen intervenir; sin embargo, este método genera un número infinito de soluciones donde las componentes demográficas muestran comportamiento armónico. Resulta entonces una tarea imposible de realizar decidir cuál de todas las posibles alternativas, todas igualmente consistentes, es la que "mejor" explica el comportamiento demográfico estudiado. La alternativa de solución que seleccione un investigador, depende de su particular apreciación del fenómeno demográfico de que se tra-

te, impidiendo que exista solución única, y aún en el caso de que en conjunto los especialistas coincidieran en que determinada solución es la mejor, no habría manera de demostrar tal situación.

Al revisar diversos trabajos de conciliación demográfica, incluso realizados por los mismos investigadores, en diferentes momentos, se observa diversidad de opiniones y supuestos, y en consecuencia de resultados. Seguramente nuestra revisión fue muy limitada, de lo contrario no sería justificada la afirmación que se hace en el texto que comentamos en el sentido de que los especialistas prácticamente han acordado la magnitud absoluta e indicadores correspondientes sobre los nacimientos, defunciones y migración neta internacional, de 1950 a 1980.

Se afirma también que ahora es posible conocer las más precisas cantidades sobre el total de habitantes del país en 1950, 1960, 1970 y 1980 y el valor aproximado de la cobertura censal en cada año. Esto último muestra inconsistencia lógica, pues si la población estimada es precisa, también lo será el nivel de cobertura que resulte.

Las cifras preliminares del censo de 1990

En el texto que comentamos se establece que "quizá el censo de 1990 presente problemas en la calidad de sus resultados definitivos, al menos en algunos temas como el de migración". Al respecto debo decir que para el diseño del cuestionario censal se realizaron numerosos seminarios y talleres de trabajo en los que participaron miembros de la academia y usuarios de la información censal. El objetivo de estas sesiones fue recoger propuestas para definir el contenido, pregunta a pregunta, del cuestionario censal. Las propuestas recibidas fueron probadas en campo y los resultados obtenidos fueron discutidos en diversas reuniones de trabajo; con este procedimiento se obtuvo la versión definitiva del cuestionario.

Cada pregunta se incluyó después de un cuidadoso análisis. En cada caso se supo lo que se ganaba y lo que se perdía en contenido de información. En particular, para el tema de migración se decidió perder la comparabilidad con censos anteriores, a cambio de tener una mayor calidad en la respuesta que ofrecen los informantes. Fue un cambio realizado evaluando pros y contras; la decisión final, apoyada en las pruebas de campo realizadas, mejora sustancialmente la calidad de la información que proporcionan los informantes, y en consecuencia se puede medir, con mayor precisión, el fenómeno de interés. Hubiera sido muy cómodo de-

jar las cosas como estaban hasta 1980, habríamos mantenido la comparabilidad, que era lo que entonces y ahora propone Rodolfo Corona, pero también habríamos mantenido la mediocre calidad de las estadísticas censales sobre la migración interna en el país.

En el texto que comentamos se afirma que de momento no es posible evaluar el censo de 1990, pero que sí es factible establecer parámetros de referencia, sobre la exactitud de sus resultados, para tal fin se realiza "un cálculo del rango donde muy probablemente se encuentre el verdadero número de todos los residentes en el país" en 1990. Al respecto se debe comentar que después de haber revisado cuidadosamente el texto que comentamos, en ninguna de sus secciones aparece, o se hace referencia, a distribución de probabilidad alguna, por lo que resulta imposible el cálculo de esa medida; es decir, la probabilidad que se menciona en el texto, no tiene relación con la que formalmente se estudia en matemáticas. Por lo anterior, sería conveniente cambiar el término probabilidad, por uno que se apegue más a la situación que el texto quiere describir; me permito sugerir, por ejemplo, el de "mucho fe".

Otro aspecto que se debe mencionar sobre este punto, es el uso del término "rango", el cual de acuerdo con el diccionario de la Real Academia de la Lengua Española, se refiere a clase, categoría o lugar que ocupa una persona en una jerarquía. Considero que hubiera sido mejor utilizar el término "intervalo" o "colección de valores".

El punto de partida para el cálculo del "rango" es la población a mediados de 1980, para el cual se supone existe una estimación confiable del total de habitantes. En forma muy rápida, en el texto que comentamos, se establecen las cifras que se juzgan más cercanas a la realidad, en cuanto al número de nacimientos y defunciones ocurridos en la década 1980-1990. Nunca se presenta prueba alguna sobre la calidad de las estimaciones que se consideran más adecuadas. Tampoco se hace una descripción del método seguido para llegar a dichos valores. Aparentemente, el método utilizado combina las estadísticas generadas por el Instituto Nacional de Estadística, Geografía e Informática, las que produce el Registro Nacional de Población, las estimaciones que resultan de proyecciones de población, y sobre todo el muy particular punto de vista del autor del texto que comentamos.

Con este procedimiento se afirma que "con los elementos anteriores se pueden adoptar números plausibles sobre el crecimiento natural en México durante el decenio: 5 millones de defunciones y 24 millones de nacimientos". Por primera vez se hace uso del término "plausible" referido a cifras o números. Esto se vuel-

ve a presentar, en forma reiterada, en el resto del texto; estas cifras o números plausibles, por alguna extraña virtud, son los que el autor considera como los que más se acercan a la realidad. Al revisar el manejo que se hace de este tipo de números (los plausibles), resulta interesante observar una regla que sistemáticamente se aplica, y que nos hemos atrevido a denominar "Regla de Plausibilidad", la cual permite decidir, dentro de una colección de cifras plausibles, cuál de ellas resulta ser más plausible que las demás; para esto es necesario ordenar de menor a mayor las cifras analizadas, resultando como la más plausible aquella que ocupa el valor central.

Combinando cifras plausibles de nacimientos y defunciones, en el texto se establece que, sin incluir migración internacional, el país a mediados de 1990 tenía 86.5, 88.5 u 89.5 millones de habitantes, y aplicando la "Regla de Plausibilidad" se afirma que la cantidad más aceptable es la de 88.5 millones.

Con respecto a la migración internacional, en el texto se presentan diferentes cifras plausibles sobre el tema, y aplicando la "Regla de Plausibilidad", antes definida, se establece que la cifra más plausible es de -1.5 millones de personas.

Combinando cifras plausibles de crecimiento natural y social, en el texto se establecen tres cifras plausibles para el monto de población en 1990, seleccionando, una vez aplicada la "Regla de Plausibilidad", a 87.0 millones como la más aceptable.

En el texto se dan razones por las cuales se puede aceptar, como plausible, a -3.0 millones de personas como saldo migratorio neto para la década 1980-1990. Esta nueva cifra plausible genera tres nuevas cantidades, también plausibles, para el total de población en 1990.

Resulta extraño que el autor no haya optado por la aplicación de la "Regla de Plausibilidad", a fin de tener una sola estimación para 1990; quizá se deba a que ese compromiso resulta demasiado serio, incluso cuando se trata de cifras plausibles. Por otra parte también se debe comentar que los mismos argumentos que se presentan para hacer plausible la cifra de -3.0 millones de saldo migratorio, también pueden servir para la plausibilidad de montos mayores como -4.0, -5.0 o -6.0 millones: Sorprende que no se haya explorado esta rica gama de posibilidades.

El esfuerzo que se realiza en el texto que comentamos, establece que la subcobertura censal de 1990 puede ser del 2.3%, o del 4.0%, o del 4.6%, o del 5.7%, o del 6.4%, o del 7.3%. Todos estos porcentajes se generan mediante el método de las cifras plausibles, que ingeniosamente aplica el autor.

Realizar un análisis estadístico de la calidad de las estimacio-

nes de subcobertura, que se presentan en el artículo, es una tarea imposible de realizar, ya que el procedimiento seguido para su elaboración, al no ser estadístico, impide conocer la precisión y confianza de cada una de ellas.

Sin embargo, consideramos de fundamental importancia hacer notar un error metodológico en la expresión utilizada para el cálculo de la supuesta subcobertura de 1990.

En el texto, en forma incorrecta, se calcula la supuesta subcobertura del censo de 1990 mediante la siguiente expresión

$$r = (P-C)/C$$

donde:

P es la población plausible para 1990

C es la población censal para 1990

La expresión correcta para el cálculo del indicador de cobertura en 1990 es

$$s = (R-C)/C$$

donde:

R es la población real para 1990

Mediante álgebra elemental, y tomando como base las dos igualdades anteriores es factible obtener la siguiente expresión

$$s = r - t$$

donde:

$$t = (P-R)/C$$

Resulta entonces que el indicador real de la cobertura del censo de 1990 (s), es igual al porcentaje calculado en el artículo que comentamos (r) menos el valor de t .

De lo anterior se tienen las siguientes relaciones

si $t > 0$, entonces $s < r$

si $t = 0$, entonces $s = r$

si $t < 0$, entonces $s > r$

En el documento que comentamos, el valor de t se supone igual a cero ($s = r$). Sin embargo, no se aporta ninguna evidencia que avale tal supuesto, por lo que tan posible es que $s = r$, como

que $s < r$ o que $s > r$, es decir, el porcentaje calculado no aporta ningún conocimiento objetivo sobre la cobertura del XI Censo General de Población y Vivienda de 1990.

En resumen, el indicador de cobertura calculado en el texto que comentamos, no mide el supuesto porcentaje de población que omitió el censo de 1990, y su calidad está sujeta a la veracidad, no probada, de las estadísticas de referencia utilizadas, es decir, resulta totalmente impropio calificar los resultados del levantamiento censal de 1990, con base en indicadores de calidad dudosa y que en todo caso se refieren a un concepto diferente del que intentan medir.

Los comentarios sobre el censo de 1990

En esta sección del documento se pone en duda la calidad de los resultados del censo de 1990, en función de diversas notas periodísticas, en las que se reportan problemas durante el levantamiento censal.

Manejar información periodística, en forma directa, para calificar eventos de toda índole, entre ellos el censo, puede presentar serias dificultades. Por ejemplo, imaginemos en el año de 3992 a un arqueólogo que descubre un valioso texto de nuestros días, que después de la traducción correspondiente dice, "los diablos rojos hicieron polvo a los tigres, les limpiaron la serie y en el último de los tres les colgaron nueve argollas". Sin tener más elementos informativos, el arqueólogo en cuestión procederá a elaborar sesudas hipótesis respecto al procedimiento desintegrador que habrán utilizado esos seres infernales, para hacer polvo a tres tigres, teniendo el sadismo de colgarle nueve argollas al último que desintegraron. Por supuesto que la nota en cuestión se refiere a un encuentro de beisbol y carece de trascendencia, excepto para los aficionados a ese deporte.

Seguramente que Rodolfo Corona tuvo el cuidado de documentar adecuadamente las declaraciones periodísticas que reportan supuesta falta de cobertura del censo de 1990. Es de esperarse, lo cual dudo mucho, que en su poder exista material que de manera contundente apoye las notas periodísticas en cuestión. De otra manera no entendemos que haya utilizado una revista seria como "Estudios Demográficos y Urbanos" de El Colegio de México, para exponer sesudas hipótesis sobre el tema, las cuales serían completamente equivalentes a las de nuestro desinformado e imaginario arqueólogo de 3992.

Otro punto que se trata en forma especial en esta sección, es

el que se refiere a las críticas a las cifras censales nacionales, que tienen sustento en las proyecciones de población.

Sobre este aspecto se debe mencionar que las proyecciones de población requieren, para su elaboración, llevar a cabo tres etapas básicas: determinación del año base y de la población de ese período; establecimiento de las hipótesis de comportamiento sobre nacimientos, defunciones y migración internacional; y finalmente elaborar los cálculos correspondientes. Consecuentemente, la veracidad de una proyección depende de la precisión con que se estime la población base y de que los supuestos establecidos respondan cabalmente al comportamiento real de la población. Resulta entonces que los resultados que ofrece una proyección, están sujetos a diversos errores, y que las proyecciones se llevan a cabo para aproximar la realidad, no para sustituirla.

Dar el carácter de infalibilidad a las proyecciones puede llevar a situaciones absurdas, las cuales tratamos de ilustrar con la siguiente nota: "el meteorólogo pronosticó que hoy estaría helando, y sin embargo hace un calor de los mil diablos, seguramente que el termómetro está descompuesto".

Deducir que un censo presenta problemas de cobertura, por no tener resultados equivalentes a los de una proyección poblacional, manifiesta la presencia del mismo error metodológico que ya se comentó para el cálculo del porcentaje de subcobertura; es decir, aquí también, y sin ninguna demostración, se está suponiendo que la proyección carece de errores.

Un elemento adicional que se debe considerar es que, hasta la creación del Consejo Nacional de Población (CONAPO), las proyecciones que se elaboraban para el país tenían como objetivo básico estimar el volumen de la población en el futuro con un carácter eminentemente predictivo, mientras que las que se elaboraron posteriormente más bien son normativas, es decir, pretenden medir los niveles de las variables demográficas que se deben conseguir para que se alcancen las metas de la política demográfica del país. Incluso la proyección básica del CONAPO suele llamarse "programática".

Coincidimos con el autor en que a través de los medios de difusión el INEGI ha dado a conocer, a la opinión pública, la forma en que se han realizado las diversas etapas del censo de 1990. No compartimos, para nada, su pesimismo al no encontrar ninguna acción positiva en todas las actividades realizadas durante el censo de 1990. Pesimismo que llega a su extremo cuando se sugiere que el procesamiento del censo de 1990 tendrá el mismo fin que el de 1980. Sobre este último punto, y a pesar de los malos augurios del autor, el censo mexicano se entregó el 10 de marzo de

1992, incluyendo resultados a nivel nacional, estatal, municipal, por localidad y por grupo de manzanas. Los resultados se presentan en material impreso y en diskette, acompañados de la cartografía correspondiente, impresa y digitalizada.

Comentarios finales

Como era de esperarse, el autor insiste en que los porcentajes calculados en el texto miden la subcobertura del censo de 1990, lo cual le permite establecer una "plausible hipótesis" sobre el total de población omitido.

Por los comentarios presentados antes, los números calculados por Rodolfo Corona, si bien se construyeron con la intención de medir la cobertura del censo de 1990, resultan ser indicadores de dudosa calidad, que se refieren a otro tema que no es la cobertura censal.

Debo agregar un aspecto, quizá irrelevante, pero que desde mi punto de vista muestra la precipitación con que se elaboró el artículo que comentamos. Su contenido, de acuerdo al título, se refiere a la confiabilidad de los resultados preliminares del censo de población de 1990; sin embargo, en ninguna sección se define qué es la confiabilidad de una cifra censal preliminar y mucho menos cómo se mide tal característica; el artículo se concreta a medir, supuestamente, la cobertura censal de 1990.

Conclusiones

- * El censo de población, al poder generar montos de población y sus características, a niveles geográficos reducidos (conjunto de manzanas), se convierte en la columna vertebral de nuestro Sistema Nacional de Información Estadística.
- * Las estimaciones de cobertura censal presentadas en el artículo tienen nivel relativo, ya que dependen de la calidad de las cifras que se utilizan como contraste, sobre las cuales no se aportan elementos objetivos que aseguren su precisión y confianza.
- * El método de conciliación censal genera un número infinito de soluciones. El autor escogió aquella solución que se ajusta a la realidad que el supone impera.
- * El diseño conceptual del censo de 1990 se llevó a cabo mediante talleres y seminarios donde participaron miembros de la academia y usuarios de la información censal.
- * El procedimiento de "cifras plausibles", si bien es muy

práctico, carece de elementos que hagan factible una evaluación estadística de sus resultados, los cuales resultan ser siempre la opción menos comprometida de todas las posibles.

- * Los porcentajes de omisión censal calculados en el artículo no miden tal característica y presentan dudosa calidad.
- * La demografía merece respeto; el uso de notas periodísticas como herramienta para evaluar procesos de producción de información estadística, debe ser ajustado estrictamente al método científico.
- * Las estimaciones del total poblacional, que se llevan a cabo a través de diversos métodos, aproximan la realidad pero no la sustituyen.

El XI Censo General de Población y Vivienda de 1990, representa una acervo de información fundamental para la generación presente, y con seguridad será consultado en el futuro para entender mejor el desarrollo que tenga nuestra sociedad; consecuentemente deberá ser evaluado en forma seria y científica, de manera que los usuarios actuales y del futuro manejen en su justa dimensión los diversos materiales que ofrece.