



Fecha: agosto de 2020 (versión 2 actualizada)

Palabras clave: Censo de población y vivienda, estimaciones demográficas basadas en modelos, censo bajo cobertura.

El valor de las estimaciones demográficas basadas en modelos a la hora de planificar y preparar el censo

El censo de población y vivienda figura entre los ejercicios más complejos y masivos que una nación puede llevar a cabo en tiempos de paz. En la mayoría de los países de ingresos bajos y medianos, el censo de población y vivienda es la principal fuente de información sobre el número de personas, su distribución espacial, estructura por sexo y edad, condiciones de vida y otras características socioeconómicas clave. También ofrece una perspectiva integral sobre las viviendas y sus condiciones en el país. Los datos del censo son fundamentales para la buena gobernanza, el desarrollo, la reducción de riesgos y la respuesta a las crisis, los programas de asistencia social y los análisis de mercado. No obstante, el censo sigue siendo una tarea ardua y logísticamente cara, y lo ideal es llevarla a cabo una vez cada diez años.

Las oficinas nacionales de estadística (ONE) a menudo se enfrentan a obstáculos que pueden afectar al éxito de la implementación de un censo nacional de población, a la calidad de sus datos y a su finalización. Se están desarrollando y evaluando enfoques estadísticos innovadores para apoyar y, posiblemente, mejorar la planificación e implementación de los censos. Las estimaciones demográficas basadas en modelos, con sus estimaciones de densidad poblacional para áreas pequeñas, ofrecen la posibilidad de apoyar la planificación de los censos y mejorar su calidad.

Los enfoques de modelización estadística para la estimación de poblaciones se desarrollan en base a la correlación entre densidad demográfica y capas de datos con covariables geoespaciales. Estas relaciones correlativas pueden incorporarse a un modelo estadístico para predecir la densidad demográfica, con ciertos niveles de incertidumbre, dependiendo de los datos de entrada iniciales. Estos modelos ofrecen la ventaja de un coste relativamente bajo y pueden ofrecer formas alternativas de obtener estimaciones demográficas más recientes que sirvan de base para la planificación e implementación de los censos. Pero, al tratarse de probabilidades, las estimaciones basadas en modelos no deben tratarse de la misma forma que los censos de población reales, sino como una oportunidad de perfeccionar la planificación cuando el censo más reciente se encuentre en gran medida desactualizado. Concretamente, las estimaciones dependen de muestras de encuestas de hogares existentes o realizadas específicamente, tales como las que se generan al realizar enumeraciones domiciliarias para las encuestas. Estas muestras se combinan con capas de datos con covariables geoespaciales para crear la base del modelo estadístico. Una vez se ha creado el modelo, pueden emplearse otras fuentes de datos independientes, como las encuestas de hogares, para verificar las nuevas estimaciones demográficas basadas en modelos.

La población se modeliza como una función de capas de datos geoespaciales basadas en imágenes de satélite u obtenidas de datos geoespaciales, tales como la distancia a las carreteras, la intensidad del alumbrado nocturno o la clasificación de la cubierta terrestre. A menudo, los mejores indicadores de densidad



Nota de Orientación Técnica

demográfica son la presencia, densidad y patrones de estructuras construidas. Asumiendo que las poblaciones normalmente se localizan en zonas de asentamientos humanos, podemos estimar la población de un país combinando las densidades demográficas con mapas de zonas de asentamientos. Esto requiere que se lleve a cabo una cartografía de asentamientos, normalmente utilizando imágenes recientes de satélite de muy alta resolución. Es entonces cuando el modelo puede predecir densidades demográficas o personas por unidad de superficie para cualquier unidad territorial, como las zonas administrativas o las áreas censales, o incluso las cuadrículas de 100 metros por 100 metros.

Con respecto a la fiabilidad, las estimaciones basadas en modelos contienen un grado de incertidumbre que puede resultar en intervalos de confianza en orden de magnitud con respecto a la densidad demográfica estimada. Hasta la fecha, no existe ningún mecanismo para determinar el nivel de incertidumbre aceptable, o si debería tenerse en cuenta a la hora de tomar decisiones. Para indicadores generales basados en datos de encuestas, la recomendación es considerar que un indicador no es robusto cuando la desviación estándar relativa (desviación típica por encima de la media) sea inferior al 33 por ciento. Aún así, incluso cuando las estimaciones demográficas estén sujetas a imprecisiones, pueden servir como:

- aporte para la planificación del censo y la cartografía del censo
- sustituta de la enumeración en zonas inaccesibles
- conjuntos de datos operacionales comunes
- datos de entrada para actualizar el marco muestral maestro
- evaluación de la cobertura de los censos
- medio para anonimizar los resultados del censo
- combinación entre censos de población y vivienda y censos agrícolas

Aporte para la planificación del censo y la cartografía del censo

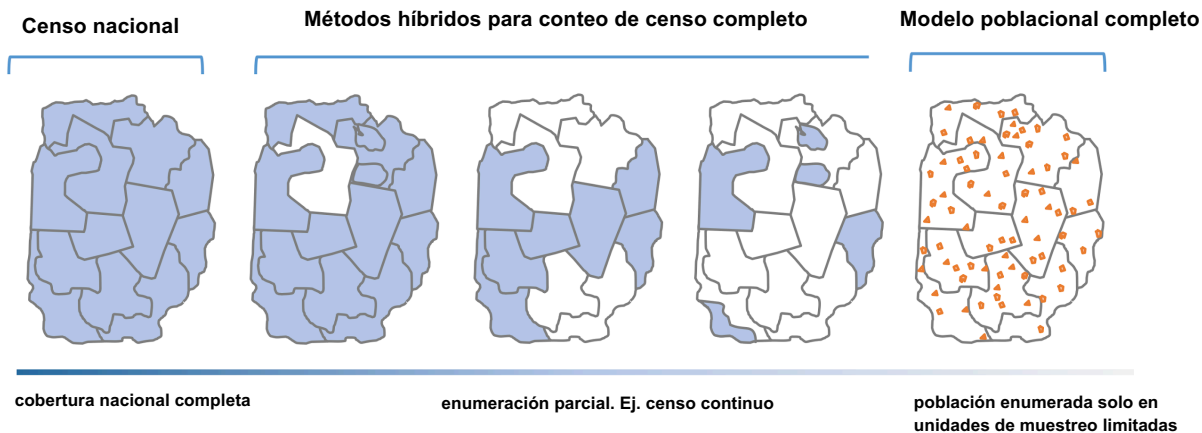
El proceso de planificación para el trabajo de campo cartográfico y censal depende fundamentalmente de las proyecciones (o estimaciones) de población por región, municipio, centro poblado y área de enumeración (AE). Estos recuentos poblacionales sirven de base para la asignación de recursos para la implementación del censo, incluyendo los fondos, el tiempo asignado para las etapas del censo, el personal y las operaciones. No obstante, las proyecciones oficiales de población elaboradas inmediatamente después del censo se pueden volver cada vez más desactualizadas a lo largo del tiempo si no se actualizan de manera constante con datos más recientes de estudios entre censos. Es más, aunque las proyecciones de población pueden ser precisas a nivel nacional y regional, también pueden resultar imprecisas a niveles más bajos, especialmente a nivel del AE. Puesto que las estimaciones demográficas basadas en modelos se basan en las covariables geoespaciales más recientes disponibles, a menudo ofrecen una estimación más precisa a nivel de distrito, centro poblado o AE.

Las estimaciones demográficas basadas en modelos, en combinación con datos geoespaciales de red de carreteras/fluvial pueden emplearse para planear las vías más eficientes para los cartógrafos. Un trabajo de campo más eficiente puede reducir los costos globales y la cantidad de tiempo necesario para completar la cartografía. Debido al enorme coste financiero de un censo, es fundamental tener en cuenta las oportunidades potenciales de ahorro. Estos ahorros de costos se pueden derivar del tiempo que un cartógrafo o

empadronador necesita para completar su trabajo de campo mediante un trazado más eficiente de las rutas, una planificación más precisa del número de personal necesario para cada AE, etc. Es más, el mapeo de asentamientos basado en imágenes de satélite mejora la planificación y actualización de la cartografía ya que puede identificar pequeños asentamientos que anteriormente no aparecían en los mapas. Además, las estimaciones demográficas basadas en modelos, junto con los datos sobre redes de carreteras y redes fluviales, pueden utilizarse para obtener nuevas áreas de enumeración.

Sustituto de la enumeración en zonas inaccesibles

Al realizar un censo, algunos países se enfrentan a problemas de seguridad y acceso en ciertas partes de su territorio, sin posibilidad de enviar encuestadores al lugar. En esos casos, es posible crear modelos de las cifras y densidades demográficas de zonas inaccesibles, utilizando bien los datos del censo recopilados, bien datos de encuestas especialmente diseñadas (Figura 1). La exactitud de las estimaciones dependerá de la calidad de la enumeración en las zonas accesibles, las propiedades intrínsecas del modelo y la probabilidad de las hipótesis de modelización. El modelo de población desarrollado siempre varía según el contexto específico (es decir, disponibilidad de datos y esfera de la población) con el fin de producir las mejores estimaciones demográficas posibles en base a los datos y tiempo de entrada disponibles.



Conjuntos de datos operacionales comunes

Los conjuntos de datos operacionales comunes (COD, por sus siglas en inglés [Common Operational Datasets]) son conjuntos de datos de referencia necesarios para apoyar las operaciones y toma de decisiones de todos los actores en las respuestas humanitarias. Los COD-PS (Conjuntos de datos operacionales sobre estadísticas de población [Common Operational Datasets on Population Statistics]) son conjuntos de datos de población actualizados desglosados por sexo y edad en el nivel geográfico más bajo disponible. Deben emplearse en asistencia humanitaria únicamente para responder a necesidades de datos en respuestas e intervenciones en situaciones de crisis. Así, su principio rector son los criterios más flexibles y pragmáticos de la «mejor opción disponible» en lugar de las normas estadísticas oficiales. En dicho caso, es posible estimar COD-PS y densidades de zonas inaccesibles utilizando los datos del censo parcialmente recopilados para construir un modelo. De esta forma, cuando los datos demográficos basados en modelos son la mejor fuente

disponible para obtener información actualizada desglosada por sexo, edad y zona geográfica, pueden emplearse como COD-PS para apoyar las operaciones y toma de decisiones en situaciones de asistencia humanitaria.

Marcos muestrales maestros y estimaciones demográficas

Una vez validados los datos del censo, generalmente se elabora un marco de muestreo maestro de AE a partir de la base de datos censales. Esta muestra se utiliza durante el periodo intercensal para llevar a cabo las grandes encuestas socioeconómicas y demográficas del país. Con el tiempo, los parámetros del marco maestro suelen desviarse de los parámetros de población, debido al crecimiento poblacional y a los cambios en su distribución espacial. Dependiendo del tipo de cartografía de censo de la que se obtiene el marco muestral maestro inicial, se pueden utilizar estimaciones de población mapeada para actualizar el marco maestro de dos maneras. Si la cartografía del censo es digital, se puede actualizar el marco maestro mediante la actualización de las poblaciones de las AE a estimaciones por cuadrícula (*gridded estimates*). Si la cartografía no es digital, no es posible actualizar las poblaciones de las AE a estimaciones por cuadrícula puesto que los límites de las AE no aparecerían en formato numérico y no permitiría superponer los estratos demográficos mapeados a los de las AE. En este caso, se recomienda tomar un nuevo marco muestral maestro de la población mapeada, considerando a cada cuadrícula o conjunto de grupos contiguos de cuadrículas como unidades primarias de muestreo. Esta segunda estrategia es una mejora de los muestreos de la zona conocida puesto que, en este caso, contamos con una estimación más exacta de la población del «área» (que, en este caso, es una malla o un grupo contiguo de mallas). Con respecto a otros usos, esto solo es posible cuando las estimaciones basadas en modelos sean lo suficientemente sólidas como para reflejar cambios en los parámetros demográficos.

Evaluación de la cobertura de los censos

Con carácter general, tras la fase de enumeración, normalmente se evalúa la cobertura del censo a través de diferentes métodos y fuentes de datos (proyecciones demográficas, datos del registro civil, datos sanitarios, datos sobre educación, etc.), que pueden incluir la implementación de una encuesta post-censal.

Cuando se trata de estimaciones demográficas geoespaciales, tenemos a nuestra disposición una fuente adicional de comparación. Para que las estimaciones demográficas geoespaciales sirvan como fuente para la evaluación de la cobertura del censo, la estimación debe realizarse mediante el uso de datos de áreas de enumeración con una cobertura perfecta y una buena calidad de datos. Una vez más, las estimaciones geoespaciales deben tenerse en cuenta únicamente cuando el rendimiento de los modelos sea alto y los supuestos en los que se fundamentan sean realistas.

Medio para anonimizar los resultados del censo

Los métodos estadísticos pueden utilizarse para anonimizar el conjunto de datos del censo (Figura 1, panel izquierdo). Se pueden agregar observaciones individuales a los totales de población por cuadrícula con alta resolución utilizando sistema de información geográfica (SIG) y métodos estadísticos. Con dichos métodos, las observaciones (por ej. totales administrativos) pueden desglosarse hasta una resolución de 100x100 m y

crear una capa de datos anónima utilizable por el gobierno para la toma de decisiones y la planificación de las encuestas y por el público en general, si se hicieran públicos.

Combinación con censos agrícolas

Combinar el censo de población y vivienda con el censo agrícola dentro de un sistema estadístico nacional resulta coherente con la Estrategia Global para el Mejoramiento de las Estadísticas Agropecuarias y Rurales¹. A nivel operacional, esta combinación puede obtenerse 1) estableciendo un marco integral de la encuesta, 2) desarrollando un marco muestral maestro para la agricultura, e 3) integrando los sistemas de gestión de datos. La experiencia demuestra que recopilar datos agrícolas limitados y bien definidos durante el censo de población y vivienda puede contribuir sustancialmente a la creación de un marco maestro eficaz para censos y encuestas agrícolas en muchos países en desarrollo. Es más, las estimaciones demográficas basadas en modelos ofrecen la oportunidad de afinar el marco muestral para censos agrícolas ya que las densidades de población reflejan la ocupación del suelo, lo que ofrece información sobre las potenciales zonas agrícolas con baja densidad demográfica y las potenciales zonas no agrícolas con alta densidad demográfica. Esta información puede utilizarse para definir con más precisión el marco muestral agrícola, mediante:

- la combinación de estimaciones basadas en modelos y la información geográfica del censo agrícola y
- la clasificación de zonas urbanas o pobladas que deben excluirse de la creación del marco muestral para encuestas agrícolas, puesto que esto ayudaría a evitar el uso de recursos para las operaciones sobre el terreno en zonas sin uso agrícola.

Conclusión

Las estimaciones demográficas basadas en modelos pueden ser útiles en diferentes etapas de un censo, desde la planificación a la implementación, desde la evaluación de la calidad hasta la utilización de datos. No obstante, es fundamental tener en cuenta que estas estimaciones deben utilizarse con precaución: modelos diferentes producirán estimaciones demográficas diferentes. Afortunadamente, existen mecanismos que permiten «elegir» de forma objetiva el modelo más preciso estadísticamente y validarlo. A pesar de la capacidad de elegir el modelo intrínsecamente mejor, es imprescindible recordar que un modelo de población geoespacial no puede reflejar correctamente el mundo real. Como dijo Box, «Recuerde que todos los modelos son incorrectos; la pregunta práctica es: ¿qué nivel de incorrección deben tener para dejar de ser útiles».

La presente nota técnica (Versión2) ha sido elaborada por Sabrina Juran, Mathias Kuepie, Maureen Jones, Catalina Valle, Kathrin Weny, Mathias Nathan del Fondo de Población de las Naciones Unidas en colaboración con Andy Tatem, Attila Lazar, Douglas Leasure y Heather Chamberlain de WorldPop de la Universidad de Southampton.

¹ Organización de las Naciones Unidas para la Agricultura y la Alimentación. Estrategia global para mejorar las estadísticas agrícolas y rurales. 2010
http://www.fao.org/fileadmin/templates/ess/documents/meetings_and_workshops/ICAS5/Ag_Statistics_Strategy_Final.pdf