

Profesor Juan Carlos Serruya

Unidad de aprendizaje: Aplicaciones de las operaciones con números Naturales y Enteros.

Tema: Estadística. Muestreo

Introducción del concepto de censo, muestra y encuesta

miércoles 20 de octubre de 2010

AUGUSTO HOSZOWSKI, MATEMATICO. La ciencia y la tecnología del censo

Los censos son herramientas clave de la sociedad que muestran la realidad de un país. Cuántos somos y quiénes somos. El Jinete Hipotético se aventura aquí en los complejos caminos del Indec.



–Le pido que se presente...

–Soy matemático, especialista en muestreo. Soy profesor de Muestreo en la Universidad Nacional de Tres de Febrero y en la sede Buenos Aires de la Universidad de Bolonia. Hace 25 años que trabajo en el Instituto Nacional de Estadística y Censos (Indec), actualmente trabajo en la Dirección de Metodología.

–¿Y a qué se dedica específicamente en el Indec?

–Actualmente la tarea que tengo es, junto con un equipo de trabajo, el desarrollo de la muestra del Censo de Población, Hogares y Viviendas 2010, que será un censo con muestra. Nuestro trabajo es el diseño de esta muestra.

–¿Cuándo será el censo?

–El próximo miércoles 27 de octubre.

–¿Para qué sirve un censo de población? Se hace cada diez años, ¿no?

–El censo de población es la base que permite el diseño de las encuestas sociodemográficas: la Encuesta Permanente de Hogares, las Encuestas de Condiciones de vida, las Encuestas de Gastos, etc., se van a fundar todas en este censo de población. El censo servirá también para la planificación de las políticas públicas, para determinar el número de representantes que la población elegirá en cada jurisdicción, conocer la situación demográfica, educativa, de acceso a los servicios básicos de todos los hogares, etc., a nivel geográfico detallado. Y es una responsabilidad de todos los hogares, no solamente nosotros estamos trabajando con el máximo empeño, vamos a necesitar también de la colaboración de todos los hogares, el censo lo hacemos entre todos. Y va a servir para toda la próxima década.

–Detrás de este megaoperativo estadístico hay también un andamiaje técnico y metodológico. Usted me decía que será un censo con muestra.

–Sí, en este censo se va a aplicar una metodología que ya se había aplicado acá en la Argentina en censos anteriores, que es muy usual en todo el mundo. Se aplicó por primera vez en el censo de población de Estados Unidos de 1940. En los grandes centros urbanos, a todas las viviendas se les aplicará un cuestionario que se llama básico, y a la décima parte de las viviendas se les aplicará un cuestionario ampliado, más extenso. En el censo de octubre, los municipios que están fuera de los grandes aglomerados van a ser censados todos con el cuestionario ampliado. Se hace así por un motivo de costos, pero también porque cuanto más extenso el cuestionario, más es la carga del trabajo del censista y la carga de los respondentes. Aumenta la probabilidad de las omisiones, ya sea de viviendas o personas, el peor defecto que puede tener un censo. También porque hay ciertas variables que no hace falta medirlas a nivel geográfico muy desagregado; por ejemplo, a nivel de

radio censal, basta con una estimación a un nivel geográfico más amplio. Una muestra cumple con esto. No hay que olvidar que el objetivo primordial de un censo es contar la población y la cantidad de viviendas. Por eso decimos que será un censo con muestra. Distinto es un censo por muestra.

–Bueno, yo recuerdo los problemas que hubo en el censo de 2001, hubo barrios enteros que quedaron sin censar, fue desastroso. ¿Qué significa que un censo sea “por muestra”?

–Un censo por muestra sería distinto, por ejemplo, de como se hace desde hace poco en Francia, que fue una decisión muy polémica: se selecciona una muestra de viviendas en los grandes centros urbanos y con esa muestra se estima la población total. Un requerimiento para aplicar esto es tener un registro actualizado de casi todas las viviendas del país.

–Sin embargo, Francia es una autoridad en la materia, ¿no es así?

–En los censos modernos, lo que hace Francia es una innovación total, revolucionaria. Por ahora es el único país en aplicar esa metodología. Aunque hubo y hay quienes en Francia son críticos con esta nueva forma de “censar”. Como anécdota, antes que se aplicara esta metodología hubo en Francia presentaciones judiciales de académicos para que no se la aplicara. En realidad, en Europa, principalmente Francia, ya en el siglo XVIII la población se estimaba mediante una muestra. Se utilizaban los registros parroquiales de defunciones o nacimientos para no tener que hacer un censo exhaustivo, relacionando el número de defunciones o nacimientos con la población. Lo llamaban el “método del multiplicador”. En esa época todavía no estaba la teoría del muestreo, pero de hecho estaban haciendo muestreo sin saberlo. P.S. Laplace fue uno de los científicos que se ocupó de esto. Ya en esa época regía la obligatoriedad del registro de nacimientos y defunciones; de otra forma, esa metodología no hubiera podido aplicarse.

–Y en el marco de los cuestionamientos que pesan sobre el Indec, ¿los resultados del censo serán confiables?

–En el Indec trabajamos para que el país tenga el mejor censo. Desde la cobertura de población que tendrá, censar a todas las viviendas y hogares del país, la rigurosidad de los procedimientos con que seleccionaremos y procesaremos la muestra, mediante un programa escrito en el Insee de Francia, hasta las bases informáticas y publicaciones para que los usuarios, académicos, etc., puedan hacer uso de la información. Y algo que quiero resaltar es que los resultados de la muestra de este censo estarán disponibles en forma digital, con toda la documentación necesaria para que los analistas puedan hacer uso de ella. La oportunidad de la información censal también será cuidada, para que no suceda que los resultados estén disponibles cuando ya perdieron actualidad.

–¿Y con qué apoyos cuentan?

–Se está trabajando con un gran profesionalismo y con el máximo nivel técnico. Para el censo están trabajando las direcciones de Metodología, Demografía, Informática, los especialistas en Logística, participando no sólo el Indec sino también las direcciones provinciales de Estadística. Hay un equipo muy grande de personas trabajando a conciencia. Hay mucha gente con muchos años de experiencia en el Indec, que no improvisa. Hemos incorporado a un grupo de estudiantes de la Universidad Nacional de Tres de Febrero, quienes, además de formarse en las técnicas estadísticas más modernas, tendrán la oportunidad de participar en un el próximo censo de población. Me parece importante acercar a los estudiantes de estadística al trabajo de las estadísticas oficiales, algo que lamentablemente hace tiempo que no ocurría.

–¿Por qué cree que un organismo técnico como el Indec está tan cuestionado?

–El problema es que los institutos de estadísticas generan información muy sensible. En todo el mundo, la estadística desata controversias. El Indec cuenta desocupados, pobres, elabora índices de precios, son todos conceptos o temáticas que tienen implicancias muy fuertes, y es evidente que hay intereses que van más allá de las matemáticas, interesados en el descrédito de nuestras estadísticas y el Indec. Yo le digo, técnicamente, que los datos que se obtienen tienen la máxima confiabilidad.

–Por ejemplo, ¿cómo se elaboran las cifras de desocupación?

–El Indec mide el desempleo a partir de una muestra aleatoria de viviendas seleccionada mediante rigurosas técnicas estadísticas. Tanto en la selección de la muestra como en su procesamiento, se aplican metodologías que son modelo en el país. La mejor evidencia es que los especialistas en mercado laboral que trabajan fuera del Indec toman para sus análisis los datos y las bases de la Encuesta Permanente de Hogares, que es la encuesta que permite medir el desempleo. No toman a la encuesta del observatorio XX o la encuesta que hizo el centro de estudios YY. Las bases de la Encuesta Permanente de Hogares son públicas, están en el sitio de Internet del Indec. Los especialistas tanto de organismos públicos como privados hacen sus simulaciones, sus estudios con esa información, porque ésa es la mejor encuesta que la Argentina tiene del mercado laboral, reconocida por su altísimo nivel técnico.

–¿Cuánta ciencia hay detrás de un censo?

–La estadística en el diseño de la muestra, la depuración de la información, la informática en el

procesamiento y la cartografía digital. Pero en estadística, a diferencia de la matemática pura, trabajamos con limitaciones materiales y la realidad siempre nos presenta imponderables. Cuando se diseña una encuesta, un censo, hay que tomar en cuenta no solamente una fórmula sino también con qué recursos y limitaciones nos enfrentamos para llevar a cabo aquello que uno tiene planificado. La experiencia y el criterio de los expertos son tanto o más valiosos que la teoría pura.

–¿Qué rol ocupa la informática en un operativo como un censo?

–Estamos implementando un software escrito en el Instituto de Estadística de Canadá (Canceis), destinado a corregir inconsistencias en censos y encuestas. Una base con millones de registros y miles de variables: donde puede haber centenares de inconsistencias posibles entre las variables, no se puede corregir manualmente, ni aun en forma semiautomática. La informática permitió la aparición de técnicas automáticas de corrección o depuración de bases de datos. Una de las primeras opciones que aparecieron fue la técnica de hot-deck, que tiene la ventaja de la sencillez de programación, pero se adapta más a los casos de valores faltantes que a las inconsistencias. Para la solución de las inconsistencias, la primera solución fue quizá la llamada metodología de Fellegi-Holt. Posteriormente, en Canadá desarrollaron una nueva metodología, implementada en Canceis, escrito originalmente para la corrección de censos, pero se lo empezó a utilizar en la depuración de encuestas. Dicho en forma sencilla, se le indica al programa cuáles son las relaciones entre las variables que deben cumplirse, las llamadas “pautas de consistencia”. Canceis busca entonces la mejor forma de corregir las inconsistencias, donde se define un criterio de qué es la mejor forma. Aquí la dificultad es doble, por un lado desarrollar una metodología que encuentre, en cierto sentido, la mejor corrección; por otro lado, que demore un tiempo razonable al procesar grandes bases, digamos millones de registros, si no, es inútil. Puedo decir que junto a Brasil somos pioneros en la utilización de ese software.

Guía de actividades:

- 1.) Lee la entrevista a AUGUSTO HOSZOWSKI, MATEMATICO La ciencia y la tecnología del censo
- 2.) Identifica las palabras desconocidas, escríbelas, busca cada definición en el diccionario y cópiala.
- 3.) Responde las siguientes preguntas:
 - a) ¿Quién es AUGUSTO HOSZOWSKI?
 - b) ¿Qué es el INDEC?
 - c) ¿Para qué sirve un censo de población?
 - d) ¿Cuál es la metodología que se aplicó en el censo 2010?
 - e) ¿En qué se diferencia este censo de los anteriores?
 - f) ¿Cuál es la diferencia entre un censo con muestra y un censo por muestra?
 - g) ¿Cuáles son los requisitos para realizar un censo por muestra?
 - h) ¿Cómo se aplicará la muestra del censo en Argentina?
 - i) ¿Cuánta ciencia hay detrás de un censo?
 - j) ¿Qué rol ocupa la estadística en un censo?
 - k) Cuenta, si lo deseas, tu experiencia personal de ese día (27/10/2010).