

# **LA GUERRA DE LOS CIEN AÑOS DEL MUESTREO PROBABILÍSTICO**

*Comunicación efectuada por el Dr. Álvaro González Villalobos  
en la sesión privada extraordinaria  
de la Academia Nacional de Ciencias de Buenos Aires  
del 13 de julio de 2006*

La publicación de los trabajos de los Académicos y disertantes invitados se realiza bajo el principio de libertad académica y no implica ningún grado de adhesión por parte de otros miembros de la Academia, ni de ésta como entidad colectiva, a las ideas o puntos de vista de los autores.

## **Introducción**

Me voy a referir a algunas de mis observaciones e ideas sobre algunos campos de la estadística aplicada, especialmente a los métodos probabilísticos y la enseñanza del Muestreo Probabilístico aplicado a Censos y Encuestas Nacionales o en gran escala, es decir correspondientes a áreas extensas. Me refiero a los modelos probabilísticos aplicados para obtener las principales estadísticas nacionales que se llevan a cabo en la mayoría de los países del mundo: censos y encuestas de agricultura, población y vivienda, pesca, industria, bosques, económicas, sociales y sobre el medio ambiente. Lo que se llaman estadísticas básicas porque son aquéllas sobre las cuales se apoyan la gran mayoría de las estadísticas de un país. Contar con estadísticas confiables, oportunas y pertinentes es hoy más importante que nunca debido al auge de las economías de mercado en el mundo, ya que la competitividad asociada a la política de mercado exigen un buen uso de la información para la toma de decisiones públicas y privadas tanto a nivel nacional como internacional. Dentro de este campo, trataré en especial de los modelos probabilísticos aplicables a los países que pueden disponer de recursos limitados, o sea a la gran mayoría de los países del mundo.

El diseño de estos modelos probabilísticos en gran escala para generar las estadísticas nacionales básicas implican grandes esfuerzos: un gran número de personas, años de trabajo y recursos económicos importantes. Como el diseño de estos modelos probabilísticos constituyen una de las más importantes aplicaciones del muestreo estadístico, es necesario llevar a cabo mejoras decisivas en la enseñanza y divulgación de esos métodos.

## **La enseñanza de muestreo probabilístico**

El professor Leslie Kish, querido maestro y amigo que murió a los 90 años en el 2000 (y quien realizó durante su vida importantísimas contribuciones a la práctica y la enseñanza del Muestro Esta-

dístico) tituló a uno de sus trabajos *The Hundred Years' Wars of Survey Sampling* (1995), interesante publicación que trata de algunos aspectos del desarrollo del muestro probabilístico durante el siglo XX. El título nos ha gustado, aunque esta comunicación tenga un enfoque diferente, porque la guerra continúa. De hecho, la introducción de métodos probabilísticos en gran escala no se ha llevado a cabo en la mayoría de los países para producir las principales estadísticas nacionales. Además nos referiremos aquí principalmente al período que va desde 1933, año en el cual el genial matemático Kolmogorov consiguió lo que célebres científicos, desde Cardano en el siglo XV, no habían conseguido: incluir la teoría de probabilidades en el campo del análisis matemático. Esto permitió naturalmente a su vez dar bases sólidas a la estadística.

El hecho que hemos mencionado de que la introducción de métodos probabilísticos en gran escala no se ha llevado a cabo en la mayoría de los países para producir las principales estadísticas nacionales es sin duda una paradoja en la era de la informática y las comunicaciones (la era de John von Neumann y Claude Shannon) porque ahora hay más facilidades que nunca para poder tomar decisiones basadas en estadísticas básicas precisas, confiables y oportunas. Todos los estadísticos saben que, por definición, hay una única manera de obtener tales estadísticas nacionales precisas, confiables y oportunas, esto es, establecer un sistema nacional de estadísticas basadas en la aplicación de métodos probabilísticos.

Esta falta de interés en la mayoría de los países para promover la enseñanza y la práctica de métodos estadísticos adecuados se debe en parte a la errónea creencia que la existencia y vasta divulgación de computadoras, instrumentos computarizados, imágenes y fotos digitales satelitales, sistemas de información geográfica y gran variedad de software estadísticos, disponibles aún en los países más pobres (lo cual tiene naturalmente, enormes ventajas) pueda sustituir el conocimiento de métodos de muestreo probabilístico. Pero nada puede sustituir al conocimiento y aplicación de modelos o diseños probabilísticos adecuados, y por lo tanto para el propósito de diseñar y realizar encuestas probabilísticas los nuevos programas e instrumentos digitales por sí solos son inservibles.

Algunos pocos países no pecan de estas erróneas creencias: tal es el caso, por ejemplo en Asia, de la República de Corea y más notable aún del Japón, en donde tuve la oportunidad de escuchar a su Alteza el emperador Akihito señalar la importancia de la estadística para el Imperio. Y efectivamente su padre el emperador Hirohito, después

de la Segunda Guerra Mundial, dio el nombre de William Edwards Deming a uno de los mayores premios científicos que otorga el país. Deming, célebre estadístico de encuestas, cambió la historia del Japón al introducir y aplicar de manera masiva métodos de control de calidad. Y nosotros modestamente estamos de acuerdo con esa opinión del Emperador. Este parecer, este reconocimiennto de la importancia de la estadística aplicada es lamentablemente muy poco frecuente entre políticos y responsables directos de las políticas científicas y técnicas nacionales en la enorme mayoría de los países del mundo.

Otra razón de tal falta de interés en promover el muestreo probabilístico se puede verificar por la desaparición de importantes centros internacionales de enseñanza que contribuyeron por varias décadas a la capacitación de estadísticos aplicados y en particular a estadísticos de encuestas. Es el caso de la desaparición del CIENES, el Centro Interamericano para la Enseñanza de la Estadística, dependiente de la OEA (la Organización de los Estados Americanos), que capacitó miles de personas del continente americano durante décadas. Aún peor, el hecho que la estadística haya desaparecido del organigrama de la OEA. Otro ejemplo es el apoyo internacional que solía brindar en París el INSEE (Instituto Nacional de Estadística de Francia) para los países africanos de lengua francesa. En este caso, el centro de capacitación fue trasladado a Abidjan, Costa de Marfil, donde los estudiantes africanos perdieron tanto la experiencia de vivir en París como el conocimiento personal de conocidos estadísticos. Eso sí, pudieron apreciar de cerca la inestabilidad política africana. Además de esos ejemplos, entre otros, cabe notar que el apoyo financiero ofrecido a estudiantes extranjeros en los Estados Unidos, la Unión Europea y las organizaciones internacionales ha disminuido considerablemente. Existen aún, claro está, excelentes cursos que ofrecen algunos institutos nacionales de estadística y universidades, principalmente en los Estados Unidos.

Otra razón de la falta de interés en promover la aplicación de muestreo probabilístico en gran escala, es, en muchos países, la conveniencia de que se ignore la verdad.

### **La enseñanza de métodos de muestreo para la agricultura: El uso de métodos probabilísticos de muestreo**

Queremos en primer lugar, para mayor claridad, especificar el tipo de encuestas agropecuarias a que nos referiremos. Tomaremos

esta descripción de un resumen de nuestro libro publicado por la FAO titulado *Multiple Frame Agricultural Surveys - Current Surveys Based on Area and List Sampling Methods*, Vol. I (1996) –publicado en seis idiomas–. Se trata del diseño general de un *Programa de Encuestas sobre la Estructura y Producción Agrícola*, es decir, el diseño general de una encuesta agrícola periódica (anual o estacional), basada en métodos probabilísticos de muestreo y estimación, cuyo propósito es obtener un gran número de datos básicos oportunos y fiables sobre el sector agrícola. Se supone que se requieren estimaciones para la mayoría de las siguientes *variables*: superficies agrícolas (preparadas, sembradas y cosechadas), rindes (previstos y obtenidos), producción agrícola, inventarios de ganado, reservas de cereales, sistemas de cultivo, costos de la producción, gastos en las explotaciones y características sociales y económicas de las explotaciones agrícolas.

Los métodos descritos de muestreo con múltiples marcos de selección combinan una muestra probabilística de áreas llamadas segmentos, seleccionadas a partir de un marco de áreas, con una breve lista complementaria de explotaciones agrícolas especiales que se enumeran totalmente durante la recolección de datos sobre el terreno. Las estimaciones con múltiples marcos de muestreo combinan las estimaciones derivadas del muestreo de áreas con las estimaciones obtenidas de la lista de explotaciones agrícolas especiales.

*El marco de áreas de la encuesta:* El diseño de una muestra de áreas consiste en una muestra probabilística estratificada de *segmentos*, con un procedimiento de selección replicada. El marco de la muestra de áreas considera que el territorio se divide en varios estratos según el uso de la tierra, definidos por la proporción de tierra cultivada, el predominio de ciertos cultivos u otras características referentes al uso de la tierra. Los estratos y los segmentos de la muestra deben tener *límites físicos reconocibles y permanentes* (camino, sendas, ríos, etc.), que puedan ubicarse tanto sobre el terreno como en los materiales cartográficos utilizados para su identificación (imágenes de satélites, mosaicos de fotografías aéreas y mapas). Se utilizan muestras replicadas para evitar el cansancio de los informantes.

*El marco de lista (de explotaciones agrícolas) complementario de la encuesta:* La lista complementaria de explotaciones especiales asegura la inclusión de las explotaciones que contribuyen de manera significativa a la estimación de algunas variables importantes de la encuesta. Esta lista de explotaciones especiales puede consistir, por ejemplo, en las explotaciones con la mayor superficie total, las que dedican mayor superficie a un cultivo determinado, las que tienen

mayor número de cabezas de ganado y aves de corral, las de mayores ingresos, las de mayor número de trabajadores agrícolas, las correspondientes a una producción localizada y las que se concentran en tipos muy especializados de producción: estas categorías podrían no contar con una representación adecuada en la muestra de áreas de la encuesta, por lo tanto la precisión de las estimaciones respecto a ciertas variables podría resultar significativamente afectada. La adición de una breve lista de explotaciones agrícolas especiales a un diseño con marco de áreas presenta relativamente pocas dificultades técnicas.

*Procedimientos de recolección de datos de la encuesta:* El componente de muestreo de áreas considerado requiere una recolección anual (o estacional) de datos sobre el terreno, realizada por encuestadores que completan un cuestionario para cada *tramo* incluido en cada segmento de la muestra. Un *tramo* es la parte de una explotación (o de tierras no agrícolas) incluidas en el segmento. Los encuestadores recogen los datos para cada *tramo* mediante entrevistas personales con el productor u otra persona responsable que pueda dar información sobre el *tramo*. La recolección de datos, junto a la cumplimentación de un cuestionario, requieren a menudo la *identificación y medida de tierras agrícolas*. Para cada segmento de la muestra, el encuestador usa una fotografía aérea ampliada (o una foto satelital o un mapa o plano a escala) en donde consten los límites del segmento. El encuestador comprueba los cultivos y otros usos del terreno en cada campo (un campo es un trozo de terreno en una parcela –una fracción de terreno totalmente rodeada por otra tierra, agua, caminos, bosques, etc., que no forman parte de la explotación– que está separado del resto de la parcela por límites fácilmente reconocibles, como senderos, vallados, etc., en el cual se cultiva una variedad específica de cultivo y tiene una fecha de siembra, o una combinación de cultivos), información que también le facilita el productor de la explotación. Estas superficies agrícolas identificadas en cada segmento de la muestra pueden ser *medidas* más tarde en la oficina mediante un instrumento de medida operado con computadora.

Existen dos tipos básicos de diseños de muestreo en función de la *unidad muestral de la última etapa* y de las *reglas para asignar sus probabilidades de selección*, a saber: diseños de muestreo de áreas (donde las unidades finales de muestreo son segmentos, es decir áreas de terreno) y diseños de muestreo de lista (donde las unidades finales de selección son generalmente explotaciones agropecuarias), también llamados encuestas por muestreo con marco de áreas y encuestas por muestreo con marco de lista.

Las encuestas agrícolas con múltiples marcos son las encuestas por muestreo probabilístico que combinan más de un diseño de muestreo para obtener las estimaciones de la encuesta, combinando diseños con marcos de áreas con diseños con marcos de lista. Una encuesta agrícola con múltiples marcos incluye generalmente una muestra de áreas y muestras de listas.

### *Comparación entre diseños alternativos de encuestas agrícolas*

La elección de un diseño estadístico apropiado para una encuesta sobre la estructura y producción agrícola requiere considerar comparativamente las ventajas, las desventajas y los requisitos de los *diseños de encuesta con múltiples marcos*, los *diseños con marco de áreas* y los *diseños con marco de lista*.

### *Diseños con múltiples marcos frente a diseños con marco de lista únicamente*

A continuación se indican algunas de las ventajas, desventajas y requisitos de estos diseños con múltiples marcos para los programas de encuestas sobre la estructura y producción agrícola en comparación con los diseños basados en un marco de lista:

#### *Ventajas:*

- Errores de cobertura relacionados con la cobertura completa. El marco de áreas da una *cobertura completa* de la población de unidades de información (explotaciones o tramos). Por consiguiente, las estimaciones probabilísticas basadas en la muestra de áreas no están sesgadas por errores de cobertura. Esto no puede lograrse con una encuesta por muestreo de lista ya que, en la práctica, no puede establecerse una lista completa de explotaciones, válida en las fechas del período de recolección de datos.
- Errores de cobertura relacionados con el uso repetido del marco. Si ha de realizarse una encuesta agrícola anual o estacional, conviene saber que un marco de áreas es generalmente mucho más duradero que un marco formado por una lista de las explotaciones. Los cambios en el aprovechamiento de la tierra o en el número y situación de las explotaciones pueden reducir la precisión de las estimaciones basadas en la muestra de áreas, pero no introducen sesgos.
- Precisión de las estimaciones. Un diseño de muestra de áreas, conduce a *estimaciones más precisas de las superficies agrícolas* (va-

riable clave estudiada en todas las encuestas agrícolas) que una muestra de lista. En efecto, por definición, en el muestreo de áreas las probabilidades de selección y los factores de expansión de la muestra son proporcionales a las superficies agrícolas.

- Errores no muestrales y medición objetiva de superficies. El diseño de muestreo de áreas permite utilizar un procedimiento más adecuado y exacto para las mediciones objetivas de superficies agrícolas, lo que es una ventaja importante para llegar a estimaciones exactas de superficie. En los diseños de muestreo de lista, la medición de superficies suele hacerse únicamente en una submuestra de explotaciones durante la compilación de datos. Este procedimiento es en general lento y de aplicación engorrosa, y no es práctico cuando se trata de explotaciones formadas por parcelas muy separadas unas de otras.
- Base para encuestas sobre el rendimiento de cultivos con mediciones de cortes. En los países en desarrollo, los productores son a menudo incapaces de facilitar estimaciones fiables de las cosechas y de la producción, y existe el problema de las unidades locales de medida que con frecuencia varían de una aldea a otra o incluso de un productor a otro.
- Tamaño de la muestra. Al comparar el tamaño de la muestra de un diseño con múltiples marcos con un diseño con marco de lista puede ser necesaria una muestra de lista más grande, a causa de las varianzas entre conglomerados y dentro de conglomerados. Las probabilidades de selección de las Unidades Primarias de Muestreo (UPM) y de las etapas iniciales del diseño de muestreo de lista a menudo no pueden establecerse debidamente para que el diseño de muestreo sea eficiente.
- Cambios en el sistema de explotaciones. La construcción del marco de áreas es independiente de las explotaciones agrícolas. Por ello, el muestreo de áreas es una solución adecuada para, por ejemplo, los países o regiones donde el sistema de explotaciones y otra infraestructura agrícola han sufrido cambios profundos y estarán en situación inestable todavía durante bastante tiempo, y para los cuales el conocimiento de la producción agrícola es más importante que obtener resultados por categorías de explotaciones (por ejemplo, los países de Europa central y oriental).
- Costos de la recolección de datos. Las encuestas por muestreo de áreas son rentables considerando el costo por explotación, ya que cada segmento contiene un grupo o conglomerado de *tramos* (unidades de información). Si las unidades de una muestra de lista no

son conglomerados de explotaciones, las explotaciones que han de ser enumeradas por un encuestador podrían estar separadas por una distancia considerable, con el consiguiente aumento de los costos de la encuesta.

#### *Desventajas y requisitos*

- Requisitos cartográficos. La construcción de un marco de áreas requiere un material cartográfico sobre el cual podrá procederse a una identificación y medición exacta de las áreas. Tal variedad de materiales e instrumentos cartográficos no se requiere para las encuestas con marco de lista.
- Falta de límites físicos permanentes. Para un diseño de muestreo de áreas con segmentos dotados de límites físicos reconocibles o con segmentos cuadrados o rectangulares, la falta de límites permanentes en los mapas, imágenes de satélites y fotografías aéreas constituye un problema grave. En zonas tropicales como África occidental, por ejemplo, a causa de las condiciones climáticas y de los sistemas de agricultura migratoria, los límites cambian con frecuencia o quedan cubiertos por la maleza y no son visibles en los materiales cartográficos.
- Proximidad a la explotación del productor o del informante. Puede no ser viable ni incluso posible utilizar una muestra de áreas en algunos países por lo escabroso del terreno, la falta de fondos o a causa de ciertos hábitos sociales de la población rural. En varios países, o amplias regiones de países, los productores viven en aldeas, a menudo a cierta distancia de su explotación. En este caso, una encuesta basada en una muestra de áreas es difícil de realizar. Si se usa un diseño de muestreo de áreas, se puede modificar la estratificación para respetar las aldeas, y de tal manera que el área del segmento no se tenga en cuenta salvo para definir a los productores que residan dentro del segmento (procedimiento de estimación de segmento abierto).
- Costos de selección de la muestra. *Al principio* del programa de encuestas, los costos son más elevados para la selección de una muestra de segmentos con límites físicos reconocibles que para una muestra de lista. Sin embargo, dado que para un programa periódico de encuestas sobre la estructura y producción agrícola es preciso actualizar frecuentemente el marco de lista de las explotaciones, con objeto de controlar una cobertura adecuada y obtener estimaciones fiables, resulta fácil justificar los recursos necesarios para una muestra de áreas alternativa.

- Personal técnico. La ejecución de un programa de encuestas agrícolas sustentable basado en métodos de muestreo de áreas requiere un personal de oficina altamente calificado que esté dispuesto a soportar el tedio y a poner en los detalles la atención exacta indispensable para la construcción y el mantenimiento del marco de áreas. Se precisan estadísticos muy capacitados para el análisis de datos, para la interpretación de los resultados y los correspondientes ajustes necesarios y para el perfeccionamiento de los procedimientos de la encuesta. El marco y la muestra no serán más duraderos que el personal. Cuando es posible construir un marco de lista exacto (un marco de lista es a menudo incompleto, o sesgado o envejecido), la ventaja de un diseño con marco de lista sobre un diseño con marco de áreas se debe en general a sencillez de ejecución y a que los métodos de estimación son más sencillos, ya que suele existir una correspondencia uno-a-uno entre las unidades de muestreo y las unidades de información.

En general, en la mayoría de los países ricos, desarrollados, se han hecho grandes esfuerzos para introducir métodos probabilísticos de muestreo. En los Estados Unidos, por ejemplo, el país pionero en este campo, ya en los años 60 se introdujeron masivamente los métodos probabilísticos de muestreo (en sustitución de los procedimientos no probabilísticos) para las estadísticas básicas de manera de ofrecer datos más confiables en su programa de estimaciones agropecuarias nacionales.

*Importantes métodos estadísticos de muestreo no incluidos en los textos clásicos*

Métodos de muestreo para agricultura muy importantes no están incluidos en los clásicos libros sobre muestreo, que son aún hoy los de mayor divulgación y uso, a saber: el libro de Leslie Kish (1965) *Survey Sampling*, el libro de Cochran (tercera edición de 1977) *Sampling Techniques*, el clásico libro de Hansen, Hurwitz y Madow (1953) *Sample Survey Methods and Theory*, el de Deming (1960) *Sample Design in Business Research*, el de Murthy (1967) *Sampling Theory and Methods*, Des Raj (1972) *The Design of Sample Surveys*, el libro de Desabie (1966) muy usado en países de lengua francesa *Théorie et Pratique des Sondages*, el de P. V. Sukhatme & B. V. Sukhatme (1970) *Sampling Theory of Surveys with Applications*, y finalmente el libro de Yates (cuarta edición del 1981) *Sampling*

*methods for censuses and surveys*. Inclusive conocidos libros sobre estadísticas agropecuarias no incluyen algunos de los más importantes métodos de muestreo ya aplicados en los Estados Unidos y en otros países del mundo. Por ejemplo los libros de S. S. Zarkovich (1965) *Sampling Methods and Censuses* o el de Kish (1989) *Sampling Methods for Agricultural Surveys*, ambos publicados por la FAO. Lo mismo vale para libros más recientes pero que no tienen la divulgación de los clásicos que hemos mencionado.

Esto se debe a varias razones. Una muy importante, a mi entender, es una razón histórica: el hecho que en los Estados Unidos las estadísticas agropecuarias eran producidas por agencias diferentes del Gobierno. Y también el estilo de los destacados estadísticos responsables en las diferentes agencias, universidades y otros centros de investigación de los Estados Unidos. Como resultado de estas rivalidades, las importantísimas investigaciones y aplicaciones llevadas a cabo en el Departamento de Agricultura (USDA) no fueron coordinadas con las investigaciones y las aplicaciones realizadas por la Oficina de Censos de los Estados Unidos (BUCEN) y otros centros de muestreo nacionales y privados y con las universidades. Los estadísticos ligados a estas últimas instituciones fueron los que escribieron los libros hoy clásicos de los cuales la mayor parte de los estudiantes estudian muestreo estadístico, aun cuando no incluyen, como hemos ya indicado, métodos de muestreo fundamentales para la agricultura que además son ampliamente utilizados alrededor del mundo.

En el “Bureau of the Census” (BUCEN), por ejemplo, se diseñaron y ejecutaron las principales estadísticas básicas sobre Población y Vivienda (además de los Censos Agropecuarios, pero no las encuestas Anuales o Permanentes), bajo la dirección del eminente estadístico Morris Hansen hasta su renuncia del BUCEN cuando pasó a dirigir el Westat, centro privado de investigación de enorme prestigio. Morris Hansen y su fortísimo grupo de colegas (como por ejemplo el gran estadístico Joseph Waksberg) continuaron trabajando en el mejoramiento de los métodos de muestreo en una gran variedad de campos, y cambiaron la manera que se realiza actualmente el muestreo probabilístico.

En el Departamento de Agricultura de los Estados Unidos (USDA), por otra parte, se diseñaron y ejecutaron todas las encuestas agropecuarias excepto el Censo (realizado por el “Bureau of the Census” hasta hace algunos años, cuando tal responsabilidad fue transferida al USDA). Fue en el USDA donde conocidos estadísticos (de esa institución y contratados de diversas universidades) tuvieron a su cargo el diseño y la

ejecución de las estadísticas agropecuarias cambiando radicalmente la manera que se realiza actualmente el muestreo probabilístico en agricultura, como indicaremos en lo que sigue.

*Una breve historia de los métodos de muestro diseñados en el USDA, y en particular la introducción de muestreo de áreas y muestreo con múltiples marcos de selección*

Métodos con múltiples marcos de lista: Entre los estadísticos de encuestas que han contribuido más al mejoramiento de los métodos de muestro basados en muestro de múltiples listas cabe mencionar a Ivan Fellegi y Alan Sunter *A Theory for Record Linkage* (1969), por ejemplo. Mejoras importantes en este campo de múltiples marcos de lista fueron realizadas por estadísticos que trabajaron en otras agencias o universidades, como es el caso Margaret Gurney y María Elena González, *Estimates from Samples from Frames where Some Units Have Multiple Listings* (1972).

Métodos de muestro de áreas: Probablemente el principal método de muestreo para agricultura, completamente diferente y novedoso, fue aplicado en el diseño probabilístico de la Encuesta Anual Agropecuaria y fue el método de muestreo de áreas, no incluido en los mencionados libros clásicos de muestreo, aun cuando las investigaciones sobre este método empezaron en la década del 40. Estos nuevos métodos fueron inventados, y su aplicación fue realizada por varios estadísticos, en especial: Earl Houseman, Jessen, King y Trelogan. Cabe recordar que paralelamente a estos trabajos en los Estados Unidos, el famoso estadístico Mahalanobis introdujo el muestreo de áreas en la India.

Métodos de muestreo de múltiples marcos: Entonces otra contribución fundamental fue realizada por H. O. Hartley en los 60, quien introdujo el *muestreo de múltiples marcos*, combinando muestreo de áreas con muestreo de listas. Como consecuencia, un gran desarrollo tuvo lugar en el USDA y la Encuesta Anual Nacional Agropecuaria basada en métodos de múltiples marcos de muestreo fue implementada en todos los Estados del país. Este desarrollo se debió a distinguidos estadísticos como es el caso de Caudill, Kibler y Paul Moore en los años 60, Vogel, Bosecker, Pratt, Rockwell, Huddleston, y Hill, en los 70, Vogel, Nealon, Geuder, Fecso, Tortora, Ford, Cotter y Kott, en los 80 y 90.

Muestreo de múltiples marcos combinado con el uso de imágenes satelitales: En los años 70 empezó otra muy importante contri-

bución a las técnicas de muestreo para la agricultura, a saber el uso efectivo de imágenes de satélites para construir los marcos de muestreo de la Encuesta Nacional Agropecuaria, y también para otros propósitos como para la estimación de áreas especiales y la estimación de rindes de varios cultivos. Esta nueva contribución fue realizada principalmente por estadísticos del USDA que aprendieron técnicas de teledetección, ya que la NASA, luego de muchos años, no pudo conseguir integrar estas técnicas para usarlas con métodos de muestreo de propósitos múltiples que era lo que requería el USDA y en particular la Encuesta Agropecuaria Anual. Los principales estadísticos que introdujeron técnicas de teledetección y el uso de imágenes de satélites fueron Wigton, Bormann, Hanuschak, Geuder, Craig, Cotter y Tomczak, entre otros.

Métodos de múltiples marcos combinados con el uso de Sistemas de Información Geográfica (SIG). El CASS: La última modificación importante para la Encuesta Anual Agropecuaria de los Estados Unidos fue la gradual introducción del llamado CASS (*Computer-aided stratification and sampling*), un Sistema de Información Geográfica ad-hoc desarrollado para la construcción de los marcos de áreas (uno por Estado) y para la selección de las muestras. Esto significó una reducción drástica del personal necesario, costes y tiempo requerido, en comparación con los procedimientos tradicionales. Actualmente, todos los marcos de muestreo y la selección de las muestras en cada Estado se realizan usando el CASS.

### *Muestreo de áreas y de múltiples marcos para encuestas agropecuarias en otros países y regiones del mundo*

A lo largo de los últimos 30 años, se han aplicado modelos probabilísticos basados en muestreo de áreas o en muestro de múltiples marcos (áreas y listas de establecimientos) en docenas de países (o parte de países) e inclusive en regiones formadas por varios países, adaptando a las condiciones locales los métodos que fueron originalmente aplicados en los Estados Unidos. Además, durante los últimos años se han aplicado métodos de múltiples marcos usando Sistemas de Información Geográfica estándar y no específicos como en el caso del CASS.

Algunos de los estadísticos que han contribuido de manera significativa en el diseño probabilístico de muestras de áreas y de múltiples marcos, y en su aplicación, en países de África, América, Asia y Oceanía, son los siguientes: Bouzaffour, Carfagna, Chinnappa, El Sheikh Elbashir, Gálvez Calix, Hale, Huddleston, Khan, Kovar,

Maranda, Martins Carrilho, Moore, Nazif, Otañez, Silva, Steiner, Stepanchich, Théberge, Trépanier, Vogel, Wallace, Wigton, y el que suscribe. En Europa, se introdujeron “segmentos cuadrados” en lugar de segmentos con límites físicos identificables en el terreno, excepto en Italia donde se reconsideró y se volvieron a utilizar segmentos con límites físicos identificables. Algunos de los estadísticos de encuestas responsables por los diseños en los países europeos son Carfagna y Gallego, entre otros. Italia es el país más avanzado de Europa en muestreo probabilístico de encuestas agropecuarias.

### **Consideraciones finales y posibles soluciones**

Como consecuencia de los argumentos expuestos, y teniendo en cuenta la opinión de muchos de los colegas que he mencionado, nos parece apropiado afirmar que tal falta de capacitación sobre métodos estadísticos para encuestas agropecuarias podría ser mejorado de manera decisiva incluyendo en los planes de estudio sobre muestreo por lo menos a tres métodos que son usuales y de gran importancia: muestreo de áreas, muestreo con múltiples marcos (de áreas y listas) y muestro por puntos. Por otra parte, la enseñanza de mejores métodos de muestreo probabilístico para encuestas agropecuarias en gran escala, debe ir acompañada de capacitación en técnicas de informática, el uso de sistemas de información geográfica, técnicas de teledetección especialmente para el análisis de imágenes y fotos de satélites, el uso de instrumentos como aparatos portátiles para determinar la posición geográfica –GPSs–, instrumentos computarizados de ingreso de datos, para la transferencia de escalas, programas de análisis estadístico, etc. Y no sólo porque estas técnicas, software e instrumentos computarizados son de gran utilidad utilizados en combinación con métodos de muestreo, sino también, porque, aunque parezca mentira, dichos instrumentos están *disponibles* en la mayoría de los países, aun en países muy pobres, es decir que existen las herramientas para la aplicación en los países en desarrollo de los métodos propuestos de encuestas agrícolas con múltiples marcos de muestreo. Estos métodos pueden ser la manera más práctica para que un país produzca los datos básicos anuales que se necesitan para el sector agrícola.

Para mejorar las estadísticas agropecuarias y el conocimiento del sector agropecuario, en muchos países es necesario mejorar los métodos de muestreo, recolección de datos, uso de nuevos instrumentos

computarizados, así como nuevos métodos de análisis y divulgación de datos.

Es necesario, a través de mejores estadísticas, mejorar el conocimiento de las condiciones de vida de millones de personas que viven en la miseria.

Es necesario conocer para mejorar.

INSTITUTO DE INVESTIGACIÓN Y DESARROLLO  
de la ACADEMIA NACIONAL DE CIENCIAS DE BUENOS AIRES

MESA DIRECTIVA

- 2005-2007 -

Presidente

Dr. JULIO H. G. OLIVERA

Vicepresidente 1°

Dr. ROBERTO J. WALTON

Vicepresidente 2°

Dr. AMÍLCAR E. ARGÜELLES

Secretario

Dr. HUGO F. BAUZÁ

Prosecretario

Dr. JORGE SAHADE

Tesorero

Ing. PEDRO VICIEN

Protesorero

Dr. FAUSTO T. L. GRATTON

Director de Anales  
Académico Titular Dr. Alberto Rodríguez Galán

Consejo Asesor de Anales  
Académico Titular Dr. Amílcar E. Argüelles  
Académico Titular Dr. Mariano N. Castex  
Académico Titular Dr. Roberto J. Walton

Secretaria de Redacción  
Dra. Isabel Laura Cárdenas

Impreso durante el mes de febrero de 2007 en *Ronaldo J. Pellegrini Impresiones*,  
Bogotá 3066, Depto. 2, Ciudad Autónoma de Buenos Aires, República Argentina  
correo-e: [rjpellegrini@fibertel.com.ar](mailto:rjpellegrini@fibertel.com.ar)