

ARTÍCULOS

*La edad de oro del sondeo en Rusia (1885-1924)*¹

MARTINE MESPOULET

Maître de conférence, Université d'Angers
martine.mespoulet@wanadoo.fr

1. INTRODUCCIÓN

En la primera mitad de los años 1920, la estadística administrativa soviética combinaba ya el uso abundante del censo exhaustivo con una utilización muy diversa de las encuestas por sondeo. En la intersección entre las aportaciones del cálculo de probabilidades a la teoría de los sondeos y los interrogantes metodológicos planteados por la práctica de las encuestas, los estadísticos rusos prosiguieron con una cuestión planteada a finales del siglo XIX. Desde el conocimiento de las discusiones de sus colegas europeos, los estadísticos rusos desarrollaron una reflexión que se apoyaba en las cuestiones prácticas planteadas por las encuestas parciales realizadas entre 1885 y 1917 en los *zemstva*² de las provincias europeas del imperio ruso.

Desde los años 1890, la metodología de las encuestas por sondeo da lugar en Rusia a numerosas tentativas, que alimentaron abundantes discusiones entre los estadísticos hasta que el sondeo acabó imponiéndose al recuento exhaustivo. En el presente artículo pretendemos dar cuenta del recorrido del sondeo hacia su reconocimiento científico e institucional. El estudio del período-bisagra de los años 1885-1924 nos permitirá responder a esta cuestión, antes de que comience una nueva etapa de esta historia a finales de los años 1920, una etapa caracterizada ya por otras formas de uso institucional de los sondeos en la estadística administrativa soviética.

¹ Traducción: Alejandro Almazán.

² Creados por una ley de 1864, los *zemstva* eran asambleas territoriales de la administración local de las 34 provincias, llamadas *gubernii*, de la parte europea del Imperio ruso. En las asambleas de los *zemstva* estaban representados los propietarios individuales de tierra y los campesinos de las comunas rurales. Nombraban un consejo ejecutivo y contrataban profesionales expertos. Podría decirse con mucha prudencia, que eran una institución equivalente a las Diputaciones Provinciales españolas del siglo XIX.

Después de haber situado en su contexto institucional las primeras formas de encuestas realizadas en Rusia a finales del siglo XIX, expondremos los debates que tuvieron lugar entre los propios estadísticos acerca de la introducción de la encuesta por sondeo. La vitalidad de la reflexión de los estadísticos rusos puede ser simbolizada por el tratamiento matemático de la afijación óptima por estrato que A. J. Kovalevskii presenta en 1924³, diez años antes que J. Neyman en los Estados Unidos, concluyendo así un período que puede considerarse como la verdadera edad de oro del sondeo en Rusia.

2. UNA PRÁCTICA ESTADÍSTICA SURGIDA DE LA DEMANDA ADMINISTRATIVA

En Rusia, como en otros países, las primeras encuestas parciales se realizaron ante la imposibilidad de efectuar censos exhaustivos cada vez que una administración tenía necesidad de recoger sobre un territorio extenso informaciones necesarias para elaborar medidas de acción económica o de intervención social⁴. Aunque ya se habían realizado con anterioridad encuestas de carácter monográfico, principalmente en el marco de las expediciones etnográficas de la Sociedad Imperial Rusa de Geografía o del Estado Mayor del ejército durante los años 1840-1860, pueden considerarse como el inicio de las encuestas parciales las desarrolladas en Rusia durante los años 1870. Tales encuestas fueron realizadas por algunas oficinas provinciales del Comité Central de Estadística del Estado zarista, el TsSK (Tsentral'nyi statisticheskii Komitet), y por los estadísticos de los *zemstva* (Eliseeva, Ploshko, 1990; Svavitskii, 1961). A partir de los años 1880, el desarrollo de la técnica del sondeo adoptará siempre formas particulares, que deben ser analizadas a la luz de la configuración institucional y administrativa en las que estuvieron encuadradas.

En el origen están las comunidades tipo

En su artículo sobre el nacimiento de las encuestas por sondeo en Rusia, A. A. Gurev (1921) estima que las primeras encuestas parciales realizadas en comunidades definidas como «típicas» se remontan a los años 1870. La que fue organizada en 1875 por el comité del *oblast'*⁵ de los cosacos del Terek del TsSK surgió de la voluntad de estudiar las diferencias de situación económica y de condiciones de vida entre las distintas localidades de cosacos de esta región, pero también de la imposibilidad material de efectuar tal estudio a escala del conjunto de este territorio administrativo. Realizar una encuesta detallada y profunda

³ KOVALEVSKII publica en 1924 *Fundamentos de la teoría del método de sondeo*.

⁴ Para una presentación y un análisis de las primeras encuestas parciales realizadas en los países de Europa occidental en el siglo XIX, se puede consultar Desrosières (1993), Cap. 7.

⁵ El *oblast'* era la denominación del territorio administrativo de ciertas provincias de la Rusia del Sur, especialmente en las regiones donde habitaban los cosacos.

exigía tiempo y personal, en mayor medida de lo que podía disponer el comité local del TsSK. El gobierno propuso entonces no realizar este estudio, sino hacerlo en partes del territorio; a este fin, el comité seleccionó «comunidades típicas» que fueron escogidas en razón de sus características «medias» respecto al conjunto de las comunidades, con el objetivo de juzgar la situación de todas las otras comunidades correspondientes al tipo (*tip*) elegido, y que supuestamente se encontraban en condiciones idénticas⁶.

Al igual que el tipo de una nación es para Quetelet el «hombre medio», la «población tipo» será la que reúna más características comunes del mayor número de comunidades de un área dada. La comunidad «media» o tipo pasa así a ser considerada como la imagen de unos semejantes que, a su vez, constituyen la mayor parte de un conjunto dado.

Aunque la encuesta de Terek pertenecía al mundo de la monografía y estaba marcada por una representación del mundo dominada por la idea de «media», era muy diferente de las monografías de Le Play⁷ por varias razones esenciales. En primer lugar, aquí se trataba de monografías de comunidades enteras y no de algunas familias de la comunidad. Por otra parte, hacían el esfuerzo de seleccionar las unidades observables no de manera intuitiva sino respetando criterios precisos establecidos con anterioridad.

Por último, el ejemplo estudiado debía permitir enunciar conclusiones más generales sobre la situación media de las comunidades de un grupo-tipo dado y sobre el comportamiento medio de su población; aunque se realizaba sobre una parte, la encuesta pretendía comprender el todo. Esto planteaba dos cuestiones que iban a alimentar los debates entre los estadísticos rusos hasta los años 1920. De un lado, la división del todo en partes iba a quedar ligada de manera indisoluble a la elección de las categorías de clasificación de una población y a los criterios utilizados; de otro, la idea de media debía asociarse, de manera explícita o no, a la legitimidad de la generalización de los resultados en las encuestas parciales.

En los años 1880, se realizaron otra clase de encuestas sobre comunidades-tipo, aunque en el marco de otra institución rusa: las administraciones territoriales de gestión local de los *zemstva*. Estas instituciones, creadas en 1864, se encargaban de organizar y financiar ciertos servicios obligatorios como la previsión social y las reservas de grano, pero también intervenían en áreas como la sanidad y la enseñanza en las que gozaban de mayor libertad de acción. Enfrentadas a la necesidad de disponer rápidamente de información cuantitativa para orientar sus decisiones, las administraciones territoriales reclutaron enseguida estadísticos y multiplicaron

⁶ Ver GUREV, 1921, p. 13.

⁷ F. LE PLAY (1806-1882) es el autor de una serie de monografías de familias, algunas de las cuales han sido publicadas en *Les ouvrières européennes* (1.^a ed. En 1885, 2.^a en 1877). Las monografías de LE PLAY se caracterizan especialmente por el hecho de que las familias observadas no eran elegidas en función de su adecuación a los criterios de definición del carácter típico establecido de antemano por los propios investigadores, sino según las informaciones proporcionadas por los notables locales. Como A. DESROSIÈRES (1993) hace notar, la selección de esas familias se efectuaba principalmente a partir de redes de familiaridad.

las encuestas en el conjunto del territorio (Jonson, 1982, Mespoulet, 2001). Sin embargo, la dificultad material de organizar en cada lugar encuestas exhaustivas de forma simultánea, impuso rápidamente la realización de encuestas parciales que resolvían el problema de la elección de unidades de observación.

En la segunda mitad de los años 1880, la encuesta sobre comunidades tipo se extendió así por las oficinas de estadística de los *zemstva*, impulsadas especialmente por la sección estadística de la Sociedad Jurídica de Moscú que dirigía A.I. Chuprov⁸. Dicha Sociedad afirmó desde su creación en 1882 su voluntad de homogeneizar los programas y métodos de encuesta estadística para el conjunto del territorio ruso, aunque dicho proyecto no se concretó hasta 1887, cuando designó una comisión para elaborar los principios sobre los que podría reposar la unificación de los procedimientos de recolección de datos. En el transcurso de los trabajos llevados a cabo por la comisión durante los meses de enero y febrero, se estableció el procedimiento de organización de las encuestas parciales de los *zemstva*, delimitando siempre su utilización respecto a los censos exhaustivos. La encuesta parcial quedó entronizada como complemento, y no aún como sustituto del estudio exhaustivo del distrito rural. Su metodología estaría basada en la selección de las comunidades tipo. Realizada sobre una muestra seleccionada de forma razonada entre la población total censada en el curso de una encuesta exhaustiva previa, la encuesta debía permitir un análisis más detallado de las características económicas y sociales de un territorio concreto, y a menudo desde la óptica comparativa. Para los estadísticos de los *zemstva*, cuyo objetivo era estudiar las diferencias económicas y sociales en el ámbito rural, la operación consistía en establecer primero áreas homogéneas, para pasar después a seleccionar las comunidades tipo, a partir de las cuales era posible evidenciar las diferencias. La manera de determinar estas áreas, los criterios de clasificación y las condiciones de traslación de los resultados parciales a generales, fueron el centro de las preocupaciones y las discusiones teóricas y metodológicas de esta comunidad de estadísticos de los *zemstva*.

La conferencia de estadísticos rusos de 1887 convocada por la Sociedad Jurídica de Moscú, confirmó las recomendaciones de su comisión estadística y precisó la noción de «carácter típico»: una localidad que reunía «los rasgos económicos más importantes y las particularidades de cada área diferente» (Gurev, 1921, p. 16). El método de selección de las comunidades tipo se especificó como «una selección de las unidades de la encuesta que permita considerar los datos obtenidos como medias para su área».

Sin embargo, la idea de reemplazar un conjunto de comunidades de un área dada por una sola de ellas no se acompaña de indicaciones precisas sobre el

⁸ A. I. CHUPROV (1842-1908): Estadístico ruso del primer cuarto del siglo XIX, considerado como el padre espiritual de la estadística de los *zemstva*. Profesor de economía política y estadística en la universidad de Moscú, formó a bastantes estadísticos del *zemstvo*. Fundador de la sección de estadística de la Sociedad jurídica de Moscú, fue por otra parte miembro del Instituto Internacional de Estadística (IIS) a partir de 1885. Padre de A. A. CHUPROV.

TSCHUPROW es la forma ortográfica de su nombre más comúnmente utilizada en las obras estadísticas occidentales.

procedimiento de definición del carácter típico o sobre la cantidad de unidades a seleccionar para la observación. El debate sobre la determinación científica de la representatividad y sobre las técnicas de muestreo fue un poco más tardío, y puede situarse entre 1895 y 1925. Tanto para los estadísticos rusos, como para sus colegas europeos de esta época, las primeras encuestas parciales no cuestionaban el ideal teórico del censo exhaustivo⁹; la aspiración al estudio total de un fenómeno o de una situación se expresaba de la misma manera en la monografía, aunque en una escala humana y territorial más reducida.

Usos de la monografía

Desde la segunda mitad de los años 1880, la encuesta monográfica se convierte en el método predilecto de los estudios sobre los presupuestos familiares realizados en una parte de los *zemstva*¹⁰. La monografía de presupuestos de este período está asociada al nombre de F. A. Scherbina, director de la oficina del *zemstvo* de Voronezh durante 1884-1903, y cuyo primer estudio data de 1885¹¹. Aunque las primeras formas de estudios de presupuestos fueron realizadas por D. P. Zhuravskii¹² en 1846, fue Shcherbina quien inauguró el empleo de esta modalidad de encuesta como instrumento de observación de las condiciones de vida de las familias rurales.

A finales de los años 1880, esta clase de encuesta, presentada por A. A. Kaufman como el «tipo ruso de la monografía de presupuesto» (Kaufman, 1992), se aplicaba con mayor frecuencia al estudio de presupuestos de las explotaciones agrarias. Sin embargo, bajo el término de encuestas de nutrición, encontramos también otras formas de estudio de presupuestos sobre muestras de familias obreras de la manufactura (Lositskii, 1926). Estas encuestas se realizaron en el marco más amplio de las estadísticas sanitarias, y se centraban principalmente en la estimación de las raciones alimenticias de las familias obreras; por tanto, no constituían un objetivo de conocimiento tan vasto como el de los estudios de los presupuestos de las familias campesinas realizados según el método de Shcherbina.

Estas últimas buscaban una mejor comprensión del funcionamiento de las explotaciones agrarias y de las condiciones de vida de las familias campesinas partiendo de un análisis metódico detallado de los ingresos y gastos. Organizadas

⁹ Desde su comunicación sobre los recuentos representativos en el congreso del Instituto Internacional de Estadística de 1895, A. N. KIAER precisa: «ni que decir tiene que los recuentos parciales representativos no tienen por objetivo reemplazar a los censos generales» (DROESBEKE, TASSI, 1990, p. 46).

¹⁰ Para una comparación con la práctica de las encuestas monográficas en los países europeos, ver DESROSIÈRES (1993), Cap. 7.

¹¹ Después de pasar por la oficina del *zemstvo* de Kaluga, F. A. SCHERBINA (1849-1936) dirige el de VORONEZH. Su prolongada presencia a la cabeza de esta oficina marca la producción estadística de este *zemstvo*, especialmente en el campo de las monografías de presupuestos familiares campesinos. Estas encuestas han sido publicadas principalmente en SCHERBINA (1900). Emigrado a Praga después de 1917, muere en 1936.

¹² D. P. ZURAVSKII (1810-1856) realizó estudios de presupuestos en el marco de expediciones de «descripción de los *gubernii*» de la Academia del Estado Mayor.

sobre una muestra formada por un pequeño número de explotaciones, elegidas a partir de los datos de una encuesta exhaustiva previa, exigían cuatro días de presencia del encuestador en el hogar observado a razón de ocho o diez horas diarias. Así, entre 1887 y 1896, sólo se estudiaron 230 presupuestos en la provincia de Voronezh para una población de 235.880 explotaciones.

La monografía era una encuesta parcial de un tipo que representaba las características medias de la población de referencia, y como señalaba G. Mayr¹³, encontraba su justificación en la comparación con los métodos de observación de las ciencias naturales¹⁴. No obstante, a partir del Noveno Congreso de Naturalistas y Médicos celebrado en 1894 (Chuprov, 1894) fue objeto de una larga discusión, pues en dicho congreso, la mayoría del centenar de estadísticos presentes mostraron un neto alineamiento hacia la encuesta exhaustiva como forma de investigación. Chuprov, no obstante, preconizó la utilización de encuestas monográficas, no como complemento de la encuesta exhaustiva, sino en sustitución de la misma, dada su preocupación por observar los cambios que se producían en la agricultura de manera regular y la imposibilidad de caracterizar la evolución de las explotaciones agrarias con los datos de los censos por familias¹⁵ efectuados diez u once años atrás. Si los objetivos eran éstos, la única solución posible era realizar encuestas; la solución consistió, por tanto, en organizar encuestas a escala reducida, repetibles a intervalos regulares. Para Chuprov, «un estudio monográfico de algunas comunidades podía reemplazar, incluso de manera reducida, una encuesta exhaustiva, siempre que fuera conducida según un plan concebido de manera rigurosa y de un dispositivo estable».

¿Qué entendía Chuprov por «plan concebido de manera rigurosa»? ¿Y por «dispositivo estable»? Si las comunidades elegidas presentaban las características del conjunto estudiado, el estudio completo de una parte del todo podía arrojar alguna luz sobre el conjunto:

«[...] si conocemos una región por poco que sea, o si tenemos a mano cualquier encuesta anterior, tenemos siempre la posibilidad de elegir comunidades conformes al tipo ordinario. Las encuestas monográficas pueden tener éxito, en particular, cuando se realizan sobre la base de observaciones en masa ya realizadas. Con la condición de que la selección de los casos a estudiar se lleve a cabo con la garantía de que se tomen los que se encuentren más próximos al tipo.»

Con esto lanzaba la idea de realizar encuestas parciales sobre comunidades tipo, en sustitución de los censos exhaustivos de la provincia o el distrito rural.

¹³ G. MAYR (1841-1925): estadístico alemán cuyos trabajos eran bien conocidos en Rusia.

¹⁴ «El método de la monografía tiene que ver con la hipótesis de que la observación efectuada sobre un solo elemento, al modo como las ciencias naturales hacen las observaciones sobre plantas y minerales, es capaz de dar un conocimiento completo de todos los elementos de la misma especie» KOVALEVSKII, 1924, p. 7.

¹⁵ Los censos por familias eran censos exhaustivos de explotaciones agrícolas y de localidades de cada distrito rural de una provincia. Una primera serie de estos censos fue realizada en todas las oficinas de los *zemstva* en los años 1880. Para una descripción de su metodología, ver SVAVITSKII (1961).

Con el reemplazo del todo por la parte, la idea del carácter «típico» conduciría la estadística rusa hacia la adopción progresiva durante los primeros años 1920 de la noción de representatividad en la organización de encuestas por sondeo.

Monografía, Tipo y Media

De hecho, desde finales del siglo XIX, la monografía de una comunidad expresaba esta idea de representatividad, ya que como parte de un todo o un conjunto dado, los estadísticos que la practicaban consideraban que podía proporcionar información sobre esa totalidad de manera fiable y apta para ser utilizada. Tan sólo ponían la condición de que se respetaran estrictamente ciertas precauciones de método, especialmente, la de efectuar la selección de las unidades que se pretendía observar según un razonamiento que pertenecería al ámbito de lo que se llamará en los años 1920, «elección juiciosa» o elección razonada¹⁶. En este caso, el estudio de un pequeño número de localidades permitía la obtención de informaciones concernientes a un número mayor. Como preconizaba Chuprov, a condición de hacer la selección sobre la base de los datos de una encuesta exhaustiva anterior y con la ayuda de la media. El método de la monografía y el «tipo» aparecen así, indisociablemente ligados a la media:

«Las familias y las explotaciones típicas en una localidad dada serán aquellas cuya situación económica se caracterice, para los diferentes factores económicos, por tener valores próximos a los valores medios aritméticos deducidos de los datos recogidos con anterioridad mediante una encuesta exhaustiva de toda la localidad. La elección de la media aritmética como característica de las localidades-tipo, por una parte, y las familias-tipo en esta localidad, por otra, proporcionan una base lógica para la inferencia a toda la población de una región de los resultados obtenidos en un pequeño número de observaciones» (citado en Lositskii, 1925).

Para poder extender los resultados obtenidos al conjunto de la población basándose en las características medias de las comunidades de un distrito rural dado, bastaba que en la construcción razonada de una muestra de talla reducida se tuviera en cuenta que:

«1) Sobre la base de los datos de la estadística de un *zemstvo*, es necesario dividir un *district* en áreas, habiendo definido previamente los indicadores más importantes de la economía campesina,

2) después, hace falta encontrar una o algunas localidades que sean las más típicas en cada área.

Es importante determinar con precisión qué indicadores es necesario considerar como importantes y característicos en cada caso, ya sea en el momento

¹⁶ En 1925, el congreso de Roma del Instituto Internacional de Estadística admitirá el uso de la «elección juiciosa» junto de la elección al azar. La «elección juiciosa» puede ser definida como un método de muestreo razonado que consistía en elegir las unidades que se juzgaba poseían las características idénticas a las medias del conjunto estudiado.

de la división en áreas o en el momento de la selección de las localidades típicas» (citado en Gurev (1921), p. 19).

Sin embargo, Chuprov fue aún más lejos y en 1894 propuso utilizar las encuestas monográficas de manera autónoma, sin necesidad alguna de disponer con anterioridad de un censo exhaustivo para seleccionar las unidades de observación. Bastaba en ese caso que fuera posible dividir un distrito rural en localidades tipo, pues el carácter típico estaba determinado especialmente por las culturas, las cualidades de las tierras o las condiciones climáticas de la región estudiada. Al reconocer una forma de autonomía en la utilización de la monografía con fines heurísticos, Chuprov creaba las condiciones para la rápida difusión de esta práctica de encuestas por sondeo. El estudio de la parte por el todo se acercaba así a la legitimidad, aunque el proceso, tanto en Rusia como en los demás países europeos, estará todavía plagado de vacilaciones y apasionados debates.

3. EL SONDEO Y LOS DEBATES DE LA PROFESIÓN ESTADÍSTICA

Como testimonio de la vitalidad de las reflexiones mantenidas por los estadísticos rusos sobre sus propias prácticas de encuesta, mostraremos el debate que a propósito de la introducción de los métodos de sondeo se desarrolló en Rusia, desde los últimos años del siglo XIX, y veremos que se desarrolló al mismo tiempo y con la misma intensidad que en los demás países europeos¹⁷, incluso puede decirse que ciertos procedimientos se introdujeron con anterioridad.

El recuento exhaustivo contaba con el favor de los estadísticos de los *zemstva* que tenían como objetivo la aprehensión total de un fenómeno o una población; la estadística rusa de finales del siglo XIX se caracterizaba todavía por la preocupación de contar todo y saberlo todo. Sin embargo, desde la primera conferencia nacional de estadísticos celebrada en 1887, las dificultades materiales vinculadas a este modo de observación justificarán el recurso, en determinadas circunstancias, a procedimientos de observación menos pesados.

La construcción de muestras en las encuestas por sondeo, así como la selección de los criterios de estudio fueron objeto de numerosas discusiones en el seno de los congresos nacionales y regionales de estadísticos de los *zemstva*, ofreciendo un lugar de debate sobre las prácticas de encuestas. Los congresos sirvieron así para la divulgación de la acción de las sociedades estadísticas (Mespoulet, 2001): se presentaban y comparaban las metodologías y los resultados de las encuestas realizadas en las diferentes provincias con el objetivo de generalizar sus logros; se presentaban y recomendaban nuevos métodos de observación, pero también de tratamiento y de análisis de datos, y la evaluación de las innovaciones se realizó en congresos sucesivos. De esta manera, se fue construyendo poco a poco, en un movimiento de oscilación entre teoría y práctica, la metodología de las encuestas por sondeo.

¹⁷ Por ejemplo, KOVALEVSKII cita en su trabajo de 1924 *Fundamentos de la teoría del método de sondeo* a ARTHUR L. BOWLEY.

Las primeras tentativas de sondeo aleatorio

La doble cuestión del grado de representatividad de las encuestas monográficas y de la posibilidad de generalización que permitía fue sin duda el centro de la discusión acerca de su validez estadística. A. Desrosières (1993) ha subrayado que en Europa, a finales del siglo XIX, el debate sobre la representatividad versaba menos sobre la precisión de la medida que «sobre la posibilidad de dar cuenta de un espacio diversificado». Aunque esta constatación puede aplicarse igualmente a los estadísticos rusos, hay que señalar que algunos de ellos comenzarían a estar preocupados por la precisión de la medida. Por aquí pasaba una línea de división entre los observadores de lo social, interesados ante todo por analizar la diversidad, y los estadísticos matemáticos de las universidades, preocupados por su parte en beneficiar a la teoría estadística con las aportaciones matemáticas, en particular del cálculo de probabilidades.

En Rusia, las primeras formas de construcción de muestras que hicieron intervenir la idea del azar en la selección de las unidades observadas datan de finales de los años 1880. Según Gurev (1921), fueron obra de A. A. Kaufman¹⁸, que entre 1887 y 1890, organizó encuestas por sondeo basadas en la sustitución de la población total de referencia por una muestra, aunque de manera imperfecta según su propia declaración (Kaufman, 1922). En efecto, la construcción de muestras destinadas al estudio de la economía rural de algunos distritos de Siberia del norte no fue todavía objeto de un procedimiento de extracción al azar único y homogéneo. Por una parte, la proporción de explotaciones seleccionadas variaba entre un cuarto y un sexto según los cantones¹⁹; por otra, las muestras se construían de forma combinada, asociando el principio de la extracción mecánica²⁰ al método de determinación previo de las áreas homogéneas. De hecho, en esta época la extracción mecánica no se efectuaba estrictamente al azar, en el sentido en el que se entiende hoy. La muestra estaba formada por unidades en las que los números eran extraídos siguiendo una progresión aritmética, por ejemplo, de cinco en cinco o de diez en diez; sin embargo el primer número no se obtenía puramente al azar, puesto que en las primeras encuestas se correspondía la mayor parte de las veces con el primer número de la lista utilizada como base del sondeo y que se establecía en un orden dado, alfabético o de otro tipo.

¹⁸ A. A. KAUFMAN (1864-1919) era profesor de estadística en los Cursos Superiores para las jóvenes de San Petersburgo. Aunque universitario, se interesa mucho en las aportaciones metodológicas de la estadística de los *zemstva*.

¹⁹ Un *canton* (en plural *cantons*), subdivisión administrativa de un *district*, puede ser considerado como el equivalente a un cantón rural.

²⁰ *mekhanicheskii otbor*. He preferido traducir esta expresión de manera literal, por el de «extracción mecánica», y no utilizar la expresión «extracción sistemática», que se corresponde al uso de la expresión *mekhanicheskii otbor* en los manuales rusos contemporáneos de estadística. En efecto, a finales del siglo XIX, el *mekhanicheskii otbor*, «extracción mecánica», era asimilado a una extracción al azar por los estadísticos de los *zemstva*, pero no es seguro que se utilizara esta formulación en el mismo sentido que los estadísticos actuales, es decir, con una concepción y una representación parecidos al aleatorio. He conservado la misma traducción para el conjunto del período estudiado en este artículo.

Algunos años más tarde, en 1896, apenas dos años antes de la invención de Chuprov en el Noveno Congreso de Naturalistas y Médicos, A. V. Peshekhonov realizó, en el distrito de Kozel', de la provincia de Kaluga, la primera encuesta por sondeo aleatorio que no combinaba una selección de áreas homogéneas y de localidades tipo. A falta de la disponibilidad de datos de un censo exhaustivo previo que permitiera delimitar las áreas y caracterizar las localidades-tipo²¹, el estadístico debió buscar otro método para estudiar los presupuestos campesinos de este distrito. Decidió construir una muestra por extracción mecánica de una familia de cada diez, efectuada a partir de la lista de explotaciones de las localidades bajo el mando de cada *staroste*²². La selección se efectuó por extracción de los rangos 1, 11, 21, 31, etc., siguiendo el orden en que los nombres se presentaban en cada lista. No se recurrió a la selección razonada de casos típicos (Peshekhonov, 1898), pero la adopción de la técnica de muestreo fue más el efecto de una necesidad impuesta que una elección real: se optó por la extracción mecánica ante la imposibilidad de aplicar una selección razonada de explotaciones típicas.

Como se ve, la transición hacia el sondeo aleatorio en Rusia, en este período, fue el fruto de las sucesivas tentativas y de los diferentes experimentos realizados para dar respuesta a las dificultades con las que los estadísticos se tuvieron que enfrentar en la práctica, a la hora de construir las muestras de sus encuestas. Para Peshekhonov, la adopción de la extracción mecánica no fue el resultado de una preferencia. Como bastantes otros estadísticos de su época, continuaba concediendo una confianza mayor a la muestra de una pequeña cantidad de unidades tipo construida de forma razonada que a una muestra mayor seleccionada de forma aleatoria. La razón humana parecía más segura que las leyes del azar. Aunque esta primera forma de sondeo aleatorio fue, como suele ser el caso de una innovación metodológica, el fruto de una constrictión, más que de una voluntad real, no dejó de ser un paso adelante en la utilización de las encuestas sobre una parte del todo, cortando definitivamente el vínculo entre éstas y el censo exhaustivo. Por otra parte, no suponía una ruptura con el método de la selección de unidades tipo, sino todo lo contrario. Como explicaba el propio Peshekhonov, el procedimiento de extracción mecánica se presentaba como la posibilidad de «garantizar el carácter típico» de las explotaciones seleccionadas:

«El carácter mecánico de la selección de explotaciones, en ausencia de cualquier encuesta previa, ha sido considerado como la mejor condición, en el caso de un número importante de explotaciones observadas, para garantizar su carácter típico» (Peshekhonov, 1898, introd.).

De manera general, los años 1894-1900 se caracterizan por diversos esfuerzos sucesivos para perfeccionar las técnicas de muestreo en los *zemstva*. Sin embargo, las encuestas por sondeo quedaban restringidas a un uso complementario del censo. En ciertas provincias podían utilizarse entre dos censos con la

²¹ La oficina estadística del zemstvo de Kaluga fue creado de forma tardía en 1896.

²² El *staroste* era el responsable elegido por tres años por la asamblea de la localidad, una especie de alcalde.

finalidad particular de estudiar las transformaciones de las explotaciones agrarias en el período intermedio. Con ese objetivo comenzaron a realizarse de forma habitual encuestas por sondeo anuales (Svavitskii, 1961), aunque la mayor parte de las encuestas por sondeo se apoyaban todavía en estudios con muestras de unidades tipo. La técnica de muestreo de extracción mecánica, considerada como una extracción al azar, comenzó así a practicarse de manera combinada, con la selección de localidades tipo. Sin embargo, el paso al muestreo aleatorio no se hizo de manera inmediata; a finales de los años 1890 se permitió el juego de las leyes del azar, pero sólo ante la imposibilidad de actuar de otra manera. Lo que caracterizó la práctica de las encuestas de los estadísticos de los *zemstva* hasta los años 1910 fue el uso de una combinación de métodos y no de uno solo.

Notemos sin embargo que desde el congreso de febrero de 1898 los estadísticos de los *zemstva* habían formulado ya un esbozo de la idea de representatividad recomendando «que el número de unidades de cada tipo, destinado a un estudio detallado, sea proporcional al efectivo total de esas unidades sobre un territorio dado» (Svavitskii, 1961, pp. 57-58). En los distintos *zemstva*, se realizaron diversos intentos sucesivos de emplear técnicas de muestreo que combinaban la selección de unidades tipo y la extracción mecánica.

La búsqueda de la representatividad: la encuesta de Viatka

El ejemplo de la encuesta realizada en la provincia de Viatka entre 1900 y 1902 muestra una vez más la cantidad de métodos de encuesta por sondeo que fueron refinándose progresivamente en Rusia ante las constricciones administrativas y materiales. Al principio los estadísticos de la oficina del *zemstvo* de Viatka habían decidido organizar, en el marco de las encuestas de ámbito territorial, un censo exhaustivo de las propiedades inmuebles y las explotaciones agrícolas (Gurev, 1921, pp. 37-42). El consejo del *zemstvo* rechazó ese proyecto y no autorizó más que una encuesta sobre cada cinco localidades. Según su propia declaración Gurev, entonces director de la Oficina de Estadística, no aceptó abandonar la idea de un censo sino coaccionado y forzado; a falta de algo mejor, se resolvió entonces organizar una encuesta por sondeo. Los estadísticos de su Oficina se esforzaron en construir una muestra que a su parecer pudiera proporcionar una representación fiel del todo.

En un primer momento el territorio de la provincia fue dividido en 954 áreas construidas de forma homogénea en función de sus características naturales, históricas y económicas, según el censo exhaustivo efectuado en los años 1880. En cada área, las localidades más importantes fueron repartidas en diferentes grupos constituidos según el criterio del tamaño medio de las explotaciones agrícolas. Después, una quinta parte de las localidades fue seleccionado en los diferentes grupos según el principio de selección de localidades típicas, y además, se hizo un estudio de presupuestos sobre las explotaciones agrícolas elegidas por su carácter típico respecto a una población dada, en la proporción del 4% al 10% según los distritos. Veinte años más tarde, Gurev justificó el principio de construcción de esta encuesta de la manera siguiente:

«La encuesta parcial así construida tenía, en primer lugar, el carácter de una encuesta masiva porque abarcaba al menos el 20% de los objetos a estudiar; en segundo lugar, la regularidad del reparto de las localidades observadas en el territorio, estaba garantizada por el hecho de que eran elegidas en las áreas constituidas de manera homogénea según sus características naturales, históricas y económicas; y en tercer lugar, la selección de las localidades para el estudio de los presupuestos de las familias se realizó de manera que cada localidad seleccionada fuera la representante²³ de un grupo particular de localidades, constituido en función de características económicas comunes.» (Gurev, 1921, p. 39)

Como elemento innovador, los estadísticos de Viatka intentaron controlar la conformidad de la estructura de su muestra con la de la población del censo exhaustivo de 1880. En efecto, aunque la muestra fue construida según el método de la selección de unidades tipo, juzgaron que comprendía un número lo suficientemente grande de unidades observadas para permitir al final, una estimación de su representatividad; aunque debería decirse: «dar una representación del conjunto observado» (Gurev, 1921, p. 42), ya que el término estadístico de representatividad (*representativnost'*) no se utilizaba todavía. De hecho, lo que los estadísticos se esforzaban en evaluar era precisamente el carácter típico de las localidades y las explotaciones de la muestra. La idea de la capacidad de una muestra para proporcionar una representación fiel de un conjunto quedaba ligada a su carácter típico y esto es lo que los estadísticos de Viatka buscaban verificar. Lo hicieron comparando los valores medios y relativos obtenidos mediante la encuesta por sondeo con los datos del primer censo exhaustivo de los años 1880. Entre los indicadores empleados en la comparación utilizaron, para cada cantón rural estudiado, el número medio de hombres y de mujeres por explotación familiar, la superficie media de las tierras arables por explotación, el tamaño medio de los rebaños por explotación y el porcentaje medio de tierras cultivadas. A partir de estos datos se puso en evidencia, para cada cantón, la diferencia o la concordancia entre los resultados de la encuesta por sondeo y los del censo.

Así, a la manera del noruego Kiaer, que definió una encuesta representativa como «una exploración parcial donde la observación se hace sobre un gran número de localidades diseminadas por toda la extensión del territorio, distribuidas de manera que el conjunto de las localidades observadas forme una miniatura del territorio total»²⁴, los estadísticos de Viatka efectuaron una selección de localidades cuidadosamente repartidas por el territorio de la provincia. A sus ojos, esta característica, asociada al nutrido efectivo de la muestra, permitía considerar que esta encuesta podía «representar» el todo.

Aunque más tarde, en 1921, Gurev afirmó que la concepción de esta encuesta no había recibido la influencia de los trabajos de Kiaer (Gurev, 1921, pp. 45-48),

²³ *predstavitelem*

²⁴ Citado en DESROSIÈRES (1993), p. 279. Sobre la historia de las primeras encuestas de Kiaer, ver LIE (2002)

la proximidad de los métodos es impresionante. ¿Hay que ver aquí simplemente el testimonio de la existencia de preocupaciones paralelas y simultáneas en países diferentes, o acaso desvelar el papel jugado por las sociedades y congresos de estadística rusos como divulgadores de la circulación de las ideas, los conceptos y los métodos, entre Rusia y los otros países europeos, sin que los nombres de sus creadores les fueran sistemáticamente asociados? Es difícil imaginar que la comunicación hecha por Kiaer en el congreso del Instituto Internacional de Estadística de San Petersburgo de agosto y septiembre de 1897 no tuviera eco en Rusia. En cualquier caso, es interesante notar que, al igual que en Noruega a finales de los años 1890, los estadísticos de Viatka pusieron en funcionamiento un procedimiento de control *a posteriori* de la representatividad de una muestra construida según el método de la selección de unidades tipo combinadas con una extracción al azar.

Los debates sobre las ventajas y los inconvenientes del empleo de muestras de unidades tipo jalonan así los años 1895-1914. En su exposición ante la comisión estadística de la Sociedad Libre de Economía, en 1900, V. G. Groman²⁵ da una justificación teórica del «método típico» que no puede dejar de recordar la argumentación de Kiaer:

«El material obtenido, en cantidad suficiente y dadas las condiciones de regularidad en las observaciones para toda clase de fenómenos suficientemente numerosos, puede ser generalizado a las demás explotaciones, ya que la certidumbre con la que se puede generalizar, en el caso de este método (así como en el caso de una extracción mecánica), se sustenta en la ley de los grandes números y puede ser formulada así: si entre todos los fenómenos a estudiar observamos algunos casos tomados a título de ejemplo, con la condición de que haya una regularidad en las observaciones y un número suficientemente elevado de casos estudiados, la totalidad estará representada en miniatura²⁶» (en Gurev, 1921, p. 32).

Hay que notar que el término «miniatura» se encuentra de nuevo en la formulación empleada por Kiaer en 1897, en la sesión del Instituto Internacional de Estadística que tuvo lugar en San Petersburgo. En este año los estadísticos debatieron sobre las ventajas del «método representativo» respecto al «método tipológico» (Desrosières, 1993). En 1900, Groman desarrolló una idea que compartían la mayor parte de los estadísticos de los *zemstva* a finales de 1890: la selección de unidades tipo no era incompatible con la idea de la representatividad. Lo esencial era tener en cuenta la diversidad y la regularidad de los fenómenos observados. De hecho en 1900, se habían establecido ya los términos de un debate que se desarrollaría, tanto en Rusia como en Europa, a mediados de los años 1920: el grado de representatividad que se podía atribuir a las muestras construidas mediante selección razonada, en comparación con las construídas por selección al azar, técnica que en Rusia se asociaba todavía a la extracción

²⁵ V. G. GROMAN (1874-1940) fue uno de los principales economistas del Gosplan en los años 1920.

²⁶ v *miniature*

mecánica. La polémica²⁷ sobre las ventajas y los inconvenientes respectivos del «método representativo» y del «método tipológico», dio lugar en 1925 a una resolución del Instituto Internacional de Estadística, reunido en Roma, que admitió los dos métodos de muestreo, «la extracción al azar»²⁸ y «la selección razonada»²⁹.

En Rusia, las técnicas de muestreo combinarían la selección de unidades tipo y la extracción al azar hasta los años 1920. La construcción de sondeos en varias etapas organizados por Groman en la provincia de Penza, a principios de los años 1910, son una ilustración.

La construcción de una encuesta en varias etapas

Con el objetivo de estudiar las transformaciones de las explotaciones agrarias y las condiciones de vida en el campo, Groman imagina un protocolo de encuesta en varias etapas que integraría el sondeo en una combinación de los diferentes métodos de encuesta utilizados en Rusia durante esta época. En 1910, presenta una comunicación en el Duodécimo Congreso de Naturalistas y Médicos (Trudy, 1912) y en su exposición se esfuerza por justificar teóricamente la combinación del sondeo, la monografía y el recuento exhaustivo, en un mismo proceso de investigación:

«La encuesta por sondeo y la monografía, tomados cada uno en sí mismos, no pueden ser llamados estadísticos ya que no permiten calcular con una certeza total garantizada el número de fenómenos caracterizados por uno u otro rasgo. [...] La combinación de los métodos de la encuesta por sondeo, de la descripción monográfica y del recuento exhaustivo ofrecen, a condición de delimitar con precisión la esfera de los fenómenos a estudiar con cada método, la completa posibilidad de captar la regularidad de la vida social, así como de contar todos los fenómenos que difieren los unos de los otros respecto a las características que interesan al investigador.» (Trudy, 1912, pp. 208-209)

La opción de combinar diferentes métodos quedaba así justificada por la preocupación de captar el todo. De manera más general podemos ya intuir las tentativas metodológicas de los estadísticos en busca de una técnica de encuesta que tenga en cuenta la aportación de las herramientas matemáticas compatibles con su propia representación del mundo y la construcción del saber. Así, para Grosman, el conocimiento del todo era indisoluble del conocimiento de las diferencias que lo dividían, de las relaciones de dependencia entre los fenómenos económicos y

²⁷ Para abordar esta polémica hay que tener en cuenta la aparición de los estadísticos universitarios con formación matemática. KAUFMAN, en su manual de 1912 *Teoría y métodos de Estadística*, expresa sus reservas metodológicas hacia las monografías de presupuestos y KOVALEVSKI, en *Fundamentos de la teoría del método de sondeo*, planteó que sólo la utilización de las matemáticas puede proporcionar un cierto grado de certidumbre.

²⁸ *random sampling*

²⁹ *purpose sampling*

sociales, y de los propios factores de variación. La elección del método utilizado en cada etapa de la encuesta debía depender del tipo de factor estudiado, y a tal fin, se podían distinguir tres niveles de análisis de los fenómenos sociales: los «factores principales o indicadores-causa», los «indicadores-causa de segundo orden» y los «indicadores-efecto».

La organización de un protocolo de encuesta en niveles tenía por objeto poner en evidencia la acción de estos tres factores: el análisis de los «factores principales» debía ser del dominio de la encuesta exhaustiva, los factores de segundo orden surgían más específicamente de la encuesta por sondeo, y por último, el estudio de las consecuencias se obtenía de la monografía. Este esfuerzo en la definición del empleo de los diferentes métodos iba acompañado de indicaciones sobre la tasa de sondeo y sobre la técnica de muestreo a utilizar en la encuesta por sondeo. Para Groman, la complejidad del análisis de los fenómenos sociales justificaba teóricamente este procedimiento de encuesta a diferentes niveles, y así entre 1911 y 1913 organiza una encuesta basada en estos principios en la provincia de Penza. Entre 1914 y 1916, V. G. Shvetsovyi llevó a cabo una encuesta similar en el *oblast'* de Don.

El censo agrícola ruso de 1916 consagra también las tentativas e innovaciones en los métodos de muestreo de este período, pero en esta ocasión se organizó, por primera vez a escala de todo el territorio ruso, una encuesta por sondeo sobre el 5% de las explotaciones como complemento del censo. Por otra parte, debida las dificultades materiales de la guerra, la explotación de los datos fue efectuada por sondeo, mediante una muestra de fichas censales seleccionadas por extracción mecánica. S. S. Kon³⁰ preconizó esta solución para reducir el retraso en la publicación de los resultados (Kon, 1917)³¹. El mismo principio fue aplicado en el tratamiento de los resultados del censo agrícola de 1917. Sin embargo, harían falta todavía algunos años para que el sondeo impusiera su legitimidad científica frente al censo exhaustivo. Sólo las dificultades de la organización y el tratamiento de un gran número de encuestas desde los primeros años de funcionamiento de la administración estadística del Estado soviético pueden dar cuenta de las razones de los partidarios del estudio del todo. Las discusiones se centrarán entonces sobre las técnicas de muestreo, en particular sobre la elección entre muestra de unidades tipo y muestra aleatoria. En tal contexto, la publicación en 1924 del tratado de teoría estadística de A. G. Kovalevskii fue un suceso de primer orden.

4. A. G. KOVALEVSKII Y EL MUESTREO ALEATORIO

Poco conocido en los países occidentales y poco citado en las obras rusas de estadística, A.G. Kovalevskii es, sin embargo, uno de los estadísticos rusos más

³⁰ S. S. KON (1888-1933) realizó sus estudios en el Instituto politecnico de San Petersburgo, donde fue uno de los alumnos de A. A. CHUPROV. Durante la Primera Guerra Mundial, fue estadístico en el Ministerio de Agricultura. En 1918, dejó Rusia por Tiflis. En enero de 1921, emigró a Paris. Después a finales de 1923, partió hacia Praga.

³¹ Como haría BOWLEY en Inglaterra. Vid BOWLEY A. L. *La aplicación del muestreo a los problemas económicos y sociológicos* EMPIRIA n.º 5, 2002. Traducción y presentación J. M. ARRIBAS.

innovadores de los años 1920. Su principal contribución es un tratamiento matemático del sondeo por estratificación que precedió al propuesto por Neyman diez años más tarde. Para Droesbeke y Tassi su tratado de estadística, *Fundamentos de la Teoría del Método de Sondeo*, es a mediados de los años 1920, «el mejor ejemplo del nivel alcanzado en Rusia por los sondeos, en la práctica y la teoría» (Droesbeke, Tassi, 1990)³².

En diferentes sentidos, Kovalevskii aparece como un hombre bisagra entre la estadística rusa y la soviética. Había nacido en 1892, y no era uno de los estadísticos de los *zemstva*; al contrario, se había formado en las matemáticas más que en las encuestas sobre el terreno. En 1916, había realizado su primera tesis³³ en la facultad de matemáticas de la universidad de Kazan y a continuación se había introducido en la estadística administrativa a través del movimiento cooperativo. Su formación profesional se produjo mediante el contacto con los antiguos estadísticos del *zemstvo* de Saratov. Desde 1917 fue jefe del servicio de estadística de la Unión de Cooperativas de Consumo de la región de Saratov, y en julio de 1921 fue reclutado por la oficina local de la Dirección central de Estadística, la TsTU (Tsentral'noe Statisticheskoe Upravlenie)³⁴. A partir de estos años mantendrá una doble actividad: como jefe de la sección de demografía de la oficina estadística de Saratov, y a partir de enero de 1930, como profesor de estadística en la universidad y más tarde en el Instituto de Economía y Planificación de la localidad³⁵.

Su carrera de estadístico tiene, por tanto, una doble articulación entre la universidad y la administración estadística del Estado, pero también entre la estadística matemática teórica y la estadística administrativa. Es en este sentido un representante de la nueva estadística matemática universitaria y de la estadística administrativa de los *zemstva*. Tampoco es sorprendente que su tratado se sitúe en la encrucijada de las preocupaciones prácticas de los estadísticos de los *zemstva* relativas a las técnicas de encuesta, en especial los procesos de muestreo, y las preocupaciones teóricas de los estadísticos universitarios. Su tratamiento del muestreo en el caso de la afijación óptima por estrato lleva la marca de esta doble influencia.

Un contexto de efervescencia en la producción estadística después de 1917

Para comprender mejor la aportación de la innovación teórica y metodológica de Kovalevskii debemos situarla en el contexto científico de principios de los

³² Respecto a la historia de las encuestas por sondeo en los Estados Unidos y la aportación de J. NEYMAN, ver especialmente DIDIER (2000) y DIDIER (2002).

³³ Se trata de la tesis de candidato (*kandidatskaia dissertatsia*), que iba seguida, como hoy en día, por la de doctorado (*doktorskaia dissertatsia*) para el acceso al grado de profesor.

³⁴ La administración estadística del Estado soviético, la TsTU, fue creada por un decreto de julio de 1918, seguido en septiembre por la reglamentación relativa a la organización de sus oficinas regionales. Ver MESPOULET (2001).

³⁵ KOVALEVSKII muere en Saratov el 13 de octubre de 1933. Para conocer más datos biográficos, cf. V. P. KORNEV (1993, p. 171) y MESPOULET (1999, pp. 341-345).

años 1920, pero también en el entorno institucional y político de la estadística administrativa de la época zarista. En un artículo de 1956, S. S. Zarkovich llamaba la atención de los estadísticos occidentales sobre el hecho de que, según él, «durante la Primera Guerra Mundial y el período inmediatamente posterior [...] es precisamente en Rusia donde se realiza el trabajo más intensivo en el campo de los sondeos» (Zarkovich, 1956).

Las cuestiones planteadas por la representatividad de la parte estudiada y las condiciones de la generalización de los resultados parciales obtenidos, habían recibido ya elementos de respuesta en el contexto europeo con la comunicación de Kiaer en el congreso de Budapest de 1901, y en el ruso con la introducción del cálculo de probabilidades en los procesos de muestreo propuesta por V. I. Bortkevich³⁶ (antes de su partida hacia la universidad de Berlín a principios de siglo), por A. A. Chuprov³⁷ y por A. A. Markov³⁸. Bortkevich había propuesto recurrir al cálculo de probabilidades para comprobar la diferencia entre la estructura de la muestra y la de la población, estimando así el nivel de representatividad de los resultados obtenidos a partir de la muestra. Por su parte, A. A. Chuprov, durante el Noveno Congreso de los Naturalistas y los Médicos de 1900, había presentado una comunicación sobre los principios de la muestra aleatoria. Diez años más tarde, en 1910, hizo una exposición bastante más profunda sobre el muestreo aleatorio y la medida de la precisión de los estimadores (Chuprov, 1912). También abordó las cuestiones planteadas por la construcción de una muestra estratificada. Otros dos estadísticos, S. S. Kon y N. S. Chetverikov³⁹, ambos alumnos de A. A. Chuprov, quedaron asociados a la reflexión mantenida en Rusia sobre la teoría de los sondeos en la primera década del siglo xx.

Chetverikov fue autor de uno de los primeros artículos de fondo sobre la encuesta por sondeo publicados en Rusia después de octubre de 1917 (Chetverikov, 1919). En el texto insistió en dos ideas fundamentales: en primer lugar, la denominación de encuesta por sondeo debía reservarse únicamente a las encuestas realizadas sobre muestras construidas de manera aleatoria, según un procedimiento de extracción al azar y respetando reglas claramente establecidas; la segunda característica fundamental era que el grado de fiabilidad no dependía solamente del tamaño de la muestra, sino también de la manera en que las unidades habían sido seleccionadas. Chetverikov añadió que la precisión de la estimación dependía igualmente de la amplitud de las variaciones del fenómeno

³⁶ V. I. BORTKEVIC (1863-1931) fue un alumno de Ju. E. JANSON en la universidad de San Petersburgo. Comenzó a enseñar la estadística en esta universidad y la dejó definitivamente por la de Berlín en 1901. BORTKIEWICZ es la forma ortográfica de su nombre más utilizada en las obras de estadística occidental.

³⁷ A. A. CHUPROV (1874-1931): hijo de A. I. CHUPROV, fue profesor de estadística en el departamento de economía del Instituto politécnico de San Petersburgo. Salió al extranjero en mayo de 1917, y no retornó a Rusia después de octubre. Murió en Praga en 1926.

³⁸ A. A. MARKOV (1856-1922) había realizado estudios de matemáticas en la universidad de San Petersburgo, donde tuvo a TCHEBYCHEFF como profesor.

³⁹ N. S. CHETVERIKOV (1885-1973) estudia la teoría estadística en el Instituto Politécnico de San Petersburgo junto a CHUPROV. Dirigiría el departamento de metodología de la CSU a principios de los años 1920.

estudiado. Finalmente expuso los principios de la extracción con o sin reemplazo, de las muestras estratificadas y de la extracción mecánica.

No obstante, a principios de los años 1920, el objeto de fondo de las discusiones entre los estadísticos permanecía igual que a finales del siglo XIX: ¿cómo estudiar los fenómenos en masa con ayuda de encuestas sobre una parte del todo, sin perder el carácter informativo total del censo o de la encuesta exhaustiva, que a pesar de su atractivo interés eran más difíciles de organizar?

La enumeración total parecía aún a bastantes estadísticos rusos el método más adecuado para estudiar las características de un conjunto, así como su composición interna, las variaciones y las relaciones entre sus componentes. Una encuesta exhaustiva conducida con rigor seguía siendo sinónimo de garantía absoluta de fidelidad de los resultados respecto a la población o los hechos estudiados (Kovalevskii, 1924, p. 5). Las dificultades prácticas de organizar un recuento exhaustivo eran, no obstante, bien conocidas por todos: la escala de los programas de encuesta y el tamaño de la población, contribuían por igual a elevar el coste en recursos humanos y materiales. Estos condicionamientos prácticos, evocados a menudo desde el siglo XIX, iban acompañados, lo que es más grave, de un bajo rendimiento científico en la producción de los censos en los primeros años de la estadística soviética. Por ejemplo, el tratamiento de los censos agrícolas de 1916 y 1917 y del censo demográfico de 1920, no había terminado aún en 1924.

En el transcurso de los primeros años de la estadística del Estado soviético, la repetición a intervalos próximos de diversos censos no permitía una explotación completa, aunque el hecho más preocupante era que, cuando los resultados definitivos estaban finalmente disponibles, ya no proporcionaban una fotografía fiel de la situación observada. De esta manera no podrían proporcionar las informaciones necesarias para la planificación económica y social del nuevo Estado.

La multiplicación de las encuestas corrientes, unida a la profusión de censos, hicieron aún más patentes las dificultades materiales del funcionamiento de la administración estadística durante el período de la guerra civil, situación agravada luego por las dificultades debidas a las reducciones de efectivos impuestas a finales de 1921. Este estado de cosas favoreció indudablemente la difusión de la práctica de encuestas por sondeo.

No obstante, para un estadístico tan convencido como Kovalevskii del interés de las encuestas por sondeo, éstas tenían sus límites pues se prestaban difícilmente al estudio de los fenómenos sociales. Éstos contaban con ciertas especificidades que los hacían diferentes a los fenómenos de la naturaleza, y por tanto, no permitían pensar en hacer encuestas por sondeo indiscriminadas. La necesidad de aprehender la diversidad le parecía un obstáculo importante, pues, desde su punto de vista, la encuesta exhaustiva mantenía toda su pertinencia cuando se trataba de aprehender la diversidad. Por el contrario, el análisis de las relaciones internas en un conjunto era claramente del dominio de los sondeos. Para el estadístico, una muestra se prestaba bien a las diferentes formas de cálculo capaces de contribuir a poner en evidencia las regularidades, parecidos y diferencias entre las distintas partes de un conjunto, y también las relaciones que las caracterizaban (Kovalevskii, 1924, p. 9).

La aplicación del sondeo aleatorio chocaba entonces con el reto de analizar la diversidad social: ¿cómo conciliar, con la mayor precisión posible, un método aleatorio de selección de las unidades de la encuesta con la necesidad de trabajar sobre la diversidad de una población dada, y tener en cuenta sin embargo los subgrupos que la componen y que presentan características ya conocidas? Esta cuestión fue resuelta en 1924 por la presentación teórica que hizo Kovalevskii de la afijación óptima por estrato.

La teoría de la afijación óptima por estrato

La aportación de Kovalevskii consistió en proponer un tratamiento matemático de la idea de muestra por estratos formulada por Chuprov en 1910. Después de recordar que, en el caso de una gran diversidad interna de la masa estudiada, era necesario constituir una muestra de gran tamaño, éste formuló una solución que abriría la vía, más tarde, a los trabajos de Kovalevskii:

«Esta indicación teórica reviste también un gran interés práctico. Muestra la utilidad de los procesos de trabajo que se esfuerzan en dividir previamente el conjunto estudiado en partes más homogéneas y de organizar la investigación, según el principio del método del sondeo, en los límites de cada una de ellas por separado. Sin embargo, es necesario señalar que en esta variante del método del sondeo, la teoría no está aún suficientemente elaborada y que las condiciones en las que su utilización parecería racional no han sido completamente estudiadas.» (Chuprov, 1912)

La primera originalidad del tratado de Kovalevskii residió en el proceso de construcción de la muestra de la población de referencia, pues para estudiar una población marcada por su diversidad y por una fuerte dispersión de la característica estudiada, propone construir estratos relativamente homogéneos, de manera que resulte más precisa la estimación de los parámetros de la población de origen, a partir de los obtenidos separadamente en los diferentes estratos. Con este fin Kovalevskii utilizaba la varianza o la desviación típica de una característica, para dar cuenta de la dispersión de los valores de esa característica alrededor de la media. Cuanto menor sea la varianza, más precisa es la estimación de un parámetro. ¿Cómo hacer para reducir la varianza en una población heterogénea? Una muestra estratificada lo permite a condición de que los estratos sean lo más homogéneos posible. En ese caso ¿cómo interviene el tamaño de la muestra? En el contexto de una muestra aleatoria no estratificada, el tamaño crece a medida que se eleva la dispersión de la característica estudiada. En el caso de una muestra estratificada, la cuestión se plantea de forma diferente, porque se tiene una muestra por estrato. Si la varianza de un estrato es pequeña, lo será a su vez su muestra. No será en ese caso necesario que un estrato poco disperso sea representado por un efectivo numeroso. Lo esencial reside entonces en la elección de estratos homogéneos. Con este fin, el estadístico debía tratar de construir estratos en los que las varianzas de la característica observada fueran pequeñas. La estimación del parámetro sería

tanto más precisa cuanto menor fuese la varianza en cada estrato; el objetivo era obtener una reducción de la varianza del conjunto de la muestra. La tasa del sondeo en cada estrato se calculaba teniendo en cuenta el grado de dispersión dado por la desviación típica del estrato, pequeña cuando la dispersión era débil y grande cuando era fuerte.

La otra particularidad de Kovalevskii consistió en asociar una extracción mecánica sin reposición con una muestra estratificada obtenida de una población heterogénea, a partir de un cierto número de subconjuntos más homogéneos, llamados estratos. Para este estadístico, en el caso de una encuesta social «los resultados más precisos se obtienen precisamente mediante una extracción mecánica y no de una aleatoria» (Kovalevskii, 1924), y las fórmulas empleadas en la estimación de los errores en el caso de una extracción aleatoria pueden ser adaptados a la extracción mecánica. Kovalevskii estima entonces que la extracción mecánica sin reposición garantiza una mayor precisión en la estimación y el grado de precisión se refuerza por el empleo de una muestra estratificada⁴⁰.

Consideraba que con los estratos restablecía las condiciones que permiten proceder a una extracción mecánica, en las mismas condiciones de representatividad que las garantizadas por la extracción aleatoria. Realizando así el procedimiento del muestreo, combina el proceso de la extracción mecánica, que pertenece al campo del sondeo aleatorio, y el de la extracción razonada, que está asociado habitualmente a la selección de tipos y la monografía. Actuando con estas opciones metodológicas mantiene el vínculo, característico de la estadística rusa de esta época, entre la idea de representatividad y la noción de tipo, integrando de esta manera la praxis y la formalización matemática.

Desde este punto de vista, el trabajo de Kovalevskii aparece como una especie de síntesis entre la aplicación del cálculo de probabilidades a la teoría estadística y la aportación de las cuestiones y experiencias prácticas de los estadísticos de los *zemstva* en el terreno de la construcción de muestras. Kovalevskii formó parte de aquéllos que posibilitaron en Rusia el paso desde una estadística fuertemente articulada con la práctica del terreno a una estadística matemática, asegurando además la transición entre dos formas institucionales diferentes de producción de datos.

Sin embargo, de la misma manera que la obra iniciada por Chuprov en 1910 permaneció en barbecho hasta el tratado de 1924, estas aportaciones no dieron lugar a una aplicación rápida en la administración estadística del Estado soviético, a pesar de la gran difusión de la práctica de las encuestas por sondeo en los diferentes departamentos de la TsSU a lo largo de los años 1920. Al final de ese decenio los estadísticos de esta administración discutían todavía sobre la elección entre el procedimiento de la extracción mecánica o el de la selección de unidades tipo en el campo de la estadística agrícola; sin hacer mención de los resultados establecidos por Kovalevskii⁴¹.

⁴⁰ Para más detalle sobre la presentación teórica de KOVALEVSKII, ver MESPOULET (2000).

⁴¹ *Vestnik statistiki*, 1, 1928, pp. 174-191.

5. CONCLUSIONES

Al igual que la obra iniciada por Chuprov en 1910 es recuperada en 1924 por Kovalevskii, su tratado *Fundamentos de la teoría del método de los sondeos* no da lugar a una aplicación rápida en la Administración estadística del Estado soviético. A pesar de la gran difusión que la práctica de encuestas por sondeo tuvo en los diferentes departamentos de la TsSU a lo largo de los años 1920, al final de esta década, las discusiones de los estadísticos de esta administración trataban todavía sobre la elección del proceso de muestreo mediante la extracción mecánica o la selección de áreas y unidades típicas, especialmente en el terreno de la estadística agrícola (Lositskii, 1928). En 1936, el tratado de Kovalevskii no era mencionado todavía en el capítulo dedicado a los sondeos del manual de referencia sobre teoría estadística publicado por la TsSU (Boiarskii, 1936). El sondeo estratificado se había introducido en la práctica de la Dirección Central Estadística soviética, pero bajo una forma particular que llamaban entonces la «estratificación típica» (*tipicheskoe raionirovanie*); los estratos eran construidos según «criterios típicos», el «tipo» todavía resistía bien frente a la aleatoriedad. Aunque desde principios de los años 1920 la mayor parte de las encuestas por sondeo de las estadísticas agrícolas regulares se realizaron sobre muestras seleccionadas según el procedimiento de la extracción mecánica, el estudio de los procesos de transformación de las explotaciones agrícolas quedó como dominio privilegiado del «tipo». En los años 1930, el empleo del «tipo» fue ampliado a un nuevo contexto político, sirviendo para la construcción de los estratos según criterios de «sector económico» o de «grupo social», dando lugar a una nueva etapa de la estadística en la URSS.

BIBLIOGRAFÍA

- BOIARSKII, A. Ia., BRAND, L. S., DAVYDOVA, O. S., STAROVSKII, V. N., KHOTIMSKII, V. I., IASTREMSKII, B. S. (1936): *Statistika. Osnovy obshchei teorii (La estadística. Fundamentos de una teoría general)*, Moscú, TsSU.
- CHETVERIKOV, N. S. (1919): «O vyborochnom issledovanii» (A propósito de la encuesta por sondeo), *Vestnik statistiki*, 8-12.
- CHUPROV, A. I. (1894): «O monograficheskom opisanii otdel'nykh selenii, kak sposobe dopolnenia i podnovlenia osnovnykh zemsko-statisticheskikh issledovanii» (A propósito de la descripción monográfica de algunas aldeas como medio para completar y poner al día las encuestas estadísticas fundamentales de los *zemslva*, in *Trudy podsekcii statistiki IX s'ezda russkikh estestvoispytatelei i vrachei v Moskve 1894g.*, Chernigov.
- CHUPROV, A. A. (1912): «Vyborochnoe issledovanie» (La encuesta por sondeo), in *Trudy podsekcii statistiki na XI s'ezde estestvoispytatelei vrachei 1910g.*, Chernigov. (Publicado también en TsSU SSSR (1960), pp. 258-270.)
- DESROSIÈRES, A. (1993): *La politique des grandes nombres. Histoire de la raison statistique*, Paris, La Découverte.
- DIDIER, E. (2000): *Les sondages entre statistique et politique. Invention d'une technique d'agrégation*, thèse de doctorat, Paris, CSI.

- DIDIER, E. (2002): «Representativeness in the First United States Surveys», *Science in Context*, 3, pp. 427-445.
- DROESBEKE, J. J., TASSI, Ph. (1990): *Histoire de la statistique*, Paris, Presses Universitaires de France.
- ELISEEVA, I. I., PLOSHKO, B.G. (1990): *Istoria statistiki (Historia de la estadística)*, Moscu, Finansy i statistika.
- GUREV, A. A. (1921): «Proiskhozhdenie vyborochnogo issledovaniya i pervye ego opyty v Rossii » (El origen de la encuesta por sondeo y sus primeras tentativas en Rusia), *Vestnik statistiki*, 1-4, pp. 1-48.
- JOHNSON, R. E. (1982): «Liberal professionals and professional liberals: the zemstvo statisticians and their work», in EMMONS, T. and VUCINIĆ, W. S., *The Zemstvo in Russia. An experiment in local self-government*, Cambridge, Cambridge University Press.
- KAUFMAN, A. A. (1912, 3.º ed. 1914): *Teoria i metody statistiki (Teoría y métodos de estadística)*, Moscu, izd. I.D. Sytlna.
- KAUFMAN, A. A. (1922): *Statisticheskiala nauka v Rossii (La ciencia estadística en Rusia)*, Moscu.
- KON, S. S. (1917): *K voprosu o primenenii vyborochnogo metoda pri razrabotke sel'skokhoziaisivennykh perepisei (A propósito de la aplicación del método del sondeo en el tratamiento de los censos agrícolas)*, Petrograd.
- KORNEV, V. P. (1993): *Vidnye deiateli otechestvennoi statistiki, 1686-1990 (Los actores eminentes de la estadística nacional, 1686-1990)*, Moscu, Finansy i statistika.
- KOVALEVSKII, A. G. (1924): «Osnovy teorii vyborochnogo metoda» (*Fundamentos de la teoría del método de los sondeos*), *Uchenye zapiski Saratovskogo gosudarstvennogo universiteta*, tome II, fasc. 2, pp. 60-138.
- LIE, E. (2002): «Sampling Surveys in Norway, 1875-1906», *Science in Context*, 3, pp. 385-409.
- LOSITSKII, A. E. (1925): «Obsledovaniy pitania naseleniya» (Encuestas sobre la alimentación de la población), *Vestnik statistiki*, 10-12, pp. 85-123.
- LOSITSKII, A. (1928): «Osnovnye voprosy sel'skokhoziajstvennoi statistiki na soveshchanií mestnykh rabotnikov po sel'skokhoziajstvennoi statistike pri TsSU SSSR» (Cuestiones fundamentales de la estadística agraria examinadas en la conferencia de los trabajadores locales de la estadística agrícola de la TsSU de la URSS), *Vestnik statistiki*, 1, pp. 174-191.
- MESPOULET, M. (1999): *Personnel el production du bureau de statistique de la province de Saratov. Histoire d'une professionnalisation interrompue (1880-1930)*, Tesis doctoral, Paris, EHESS.
- MESPOULET, M. (2000): «Du tout a la partie. L'age d'or du sondage en Russie (1885-1924)», *Revue d'études comparatives Est-Ouest*, 2, pp. 5-49.
- MESPOULET M. (2001): *Statistique el révolution en Russie. Un comprornis impossible*, Rennes, Presses Universitaires de Rennes.
- PESHEKHONOV, A. V. (1898): *Statisticheskoe opisaniye Kaluzhskoi gubernii. Tom I, Kozel'skii uezd Vypusk 2. (Descripción estadística de la provincia de Kaluga. Tome I, Kozel'skii district., Fasc.2) Kaluga.*
- SHCHERBINA, F. A. (1900): *Krest 'ianskie biudzhety (Los presupuestos campesinos)*, Voronezh, izd. I.V.E.O.
- STANZIANI, A. (1991): «Statisticiens> zemstva et État dans la Russie des années 1880», *Cahiers du Monde russe et soviétique*, 4, pp. 445-468.
- SVAVITSKII, N. A. (1961): *Zemskie podvornye perepisi (Los censos de hogares de los zemstva)*, Moscu, Gosstatizdat. *Trudy podsekcii statistiki XII s «ezda russkikh estest-*

- voispytatelei i vrachei v Moskve 29dekabria 1909g. - Sjanvaria 1910g.* (Trabajos de la sección estadística del XII congreso de naturalistas y médicos), Chernigov, 1912. TsSU SSSR (1960): *Voprosy Statistiki (Cuestiones de estadística)*, Moscu, Gosstatizdat.
- ZARKOVICH, S. S. (1956): «Note on the history of sampling methods in Russia», *Journal of the Royal Statistical Society*, Serie A, 11 9, pp. 336-338.
- ZARKOVICH, S. S. (1962): «A supplement to “Note on the history of sampling methods in Russia”», *Journal of the Royal Statistical Society*, Série A, 125, pp. 580-582.

RESUMEN

Este artículo presenta las diferentes formas adoptadas por el desarrollo de las encuestas por sondeo en Rusia entre 1885 y 1924, año de la aparición del tratado teórico de A. G. Kovalevskii sobre la afijación óptima por estrato, diez años antes de la publicación de J. Neyman sobre la misma cuestión, en los Estados Unidos. Después de haber situado las formas adoptadas por las primeras encuestas en el contexto de los requerimientos de la administración en el origen, este artículo analiza la evolución de las diferentes técnicas de muestreo, ilustrado por el estudio de los diferentes tipos de aplicación en la administración rusa. Finalmente, expone los principales puntos en el tratamiento matemático de afijación óptima por estrato presentada en el tratado de A. G. Kovalevskii. Obra que puede ser considerada como el resultado de la vitalidad de la reflexión teórica mantenida por los estadísticos rusos en el campo de los sondeos desde principios del siglo XX, en la intersección de la introducción de la teoría de la probabilidad y las aportaciones de la praxis de las encuestas por sondeo de los estadísticos de los *zemstva*.

ABSTRACT

This paper shows the several ways of development of surveys in Russia between 1885 and 1924, the year of the publication of Kovalevskii's treatise on optimal sample allocation (ten years before J. Neyman's publication about the same subject in USA). Firstly, it is related the ways of adaptation of surveys to public administration requirements. Then, it is analysed the evolution of the sampling techniques, illustrated by the several types of applications used by Russian public administration. Finally, Kovalevskii's mathematical treatment of optimal sample allocation is set out. Kovalevskii's work is an outcome of the Russian statisticians' important theoretical reflections on polls since the turn of the 20th Century. That thinking is located at the intersection between the introduction of the theory of probability and the contributions of the poll praxis among the *zemstva's* statisticians.

