

LA ESTADÍSTICA EN SOCIOLOGÍA*

Maurice Halbwachs

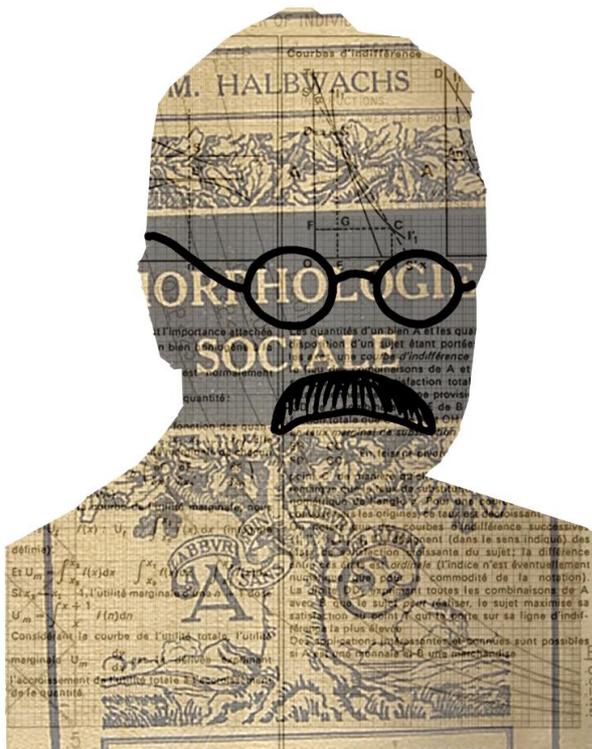


ILUSTRACIÓN: P.M.Hoyos

Mi gran amigo Simiand debía acompañarme hoy, como hace dos años, cuando hablé aquí de la ley en sociología. No será así en esta ocasión y, con enorme tristeza, me encuentro solo delante de ustedes. En mi intervención me inspiraré, fundamentalmente, en sus originales reflexiones sobre el método estadístico. Su pensamiento nos pertenece y pertenece al futuro, y si bien él ya no está aquí para guiarnos, al menos la vía que trazó permanece claramente visible. Quisiera compartir con ustedes mi convicción, reforzada con pruebas, de que no hay otra vía más directa ni más segura.

Aunque la estadística se haya aplicado progresivamente en diferentes ciencias de la naturaleza, su descubrimiento se produjo inicialmente en el dominio de las ciencias sociales. Numerosos objetos despertaron el interés inicial de los estadísticos: la población, los recursos diversos y característicos de los Estados, la esperanza de vida en diferentes grupos humanos, los nacimientos, los matrimonios, la relación entre la natalidad masculina y femenina. Por ello hay razones suficientes para pensar que tendremos una buena oportunidad de descubrir los rasgos distintivos de la investigación estadística, si observamos el uso que de ella se hace en sociología.

* Artículo original: Halbwachs, M. 1972. «La statistique en sociologie» pp. 329-348 en *Classes sociales et morphologie*. Paris: Minuit, «Le sens commun». Traducción al español de Francisco Manuel Carballo Rodríguez (2017).

A mi parecer, Simiand advirtió con acierto que la estadística aplicada a la sociología, a pesar de las similitudes cuando se trata de estudiar a grupos, no se confunde con los métodos utilizados por otras ciencias, sino que más bien los completa. La estadística es, de hecho, el único modo de reconocer las regularidades sociales.

En ocasiones se define a la estadística como la ciencia de las medias. Por ejemplo, partiendo de un cierto número de observaciones de un conjunto de individuos, extraeremos una media que nos informará sobre el índice cefálico de una raza. Es cierto que un físico también puede determinar la densidad de un cuerpo mediante una serie de observaciones particulares, a partir de las cuales calculará la media. El físico puede calcular, por ejemplo, la media de las densidades de un cierto número de barras de hierro. Sin embargo, hay una diferencia entre ambos ejemplos. En el caso de la densidad de un cuerpo, esta puede determinarse a través de una única observación si disponemos para ello del instrumento o del observador, situados en las mejores condiciones posibles. En cuanto al índice cefálico, por el contrario, no es posible obtenerlo mediante una sola observación de un único sujeto. De hecho, mediante dicho índice se define una característica que, como tal, puede no encontrarse en ninguno de los individuos observados, pero que se aplica a todo el conjunto.

La estadística, por lo tanto, se aplica a aquellos hechos que determinamos cuantitativamente mediante una selección más o menos amplia de casos particulares. Estos hechos son diferentes de los elementos individuales que los componen y no se encuentran como tales en ninguno de ellos.

En otras palabras, la estadística nos permite acceder a las características de un grupo que, si bien son reales para su conjunto, no identificaremos en ningún miembro apartado y aislado del mismo. La esperanza media de vida de un grupo de hombres, por ejemplo, no será, en general, la edad exacta alcanzada por ninguno de ellos tomados al azar, aunque sin embargo ésta sea una realidad, ya que podemos encontrarla idéntica para el mismo grupo en épocas sucesivas; también en una misma época, en distintos grupos compuestos de individuos repartidos de la misma manera en función de la edad.

Lo que sucede es que estos grupos tienen una cierta realidad, una cierta consistencia en tanto que grupos y, por lo tanto, cualquier forma de recuento no es una estadística. Un naturalista no se entretendrá en calcular el tamaño medio de los animales de un zoológico, que se encuentran agrupados en un mismo espacio por razones imputables al azar o al capricho. Algo parecido ocurre con la media de precios recogidos al azar, como los precios al por mayor o al por menor, de los salarios, de los servicios o de los alquileres. Eso tampoco es una estadística. Sí lo es, sin embargo, el cálculo del tamaño medio de los animales de

una especie, o de los precios de venta al por mayor, en un país y en un momento determinado. La misma situación la encontramos en el caso de las series de datos consideradas en el tiempo. La media de los precios recogidos con un intervalo de diez o incluso de cinco años no es una estadística, ya que no sabemos si existe alguna relación o alguna similitud entre esos años desde un punto de vista económico y qué variaciones se han producido en esos periodos. La media de precios recogidos en una serie sucesiva de años, durante un periodo de subida o de bajada, y el incremento medio anual de los precios en cada uno de dichos periodos, por el contrario, sí es una estadística, dado que las series para cada periodo representan bien a un grupo que corresponde a una realidad, que tiene una consistencia. Una agrupación en el tiempo, en el espacio, a la que exigimos un cierto orden y una cierta relación entre las partes, y estudiada como tal, es a lo que consideramos un objeto al que aplicar la estadística.

¿Es suficiente esta definición enunciada por Simiand? ¿Es aplicable a todo el conjunto o lo es también para cada una de sus partes? Se le puede objetar que, en el universo de los seres vivos, los organismos por una parte y, en su interior, los órganos por otra, constituyen también conjuntos reales y consistentes. Grupos de órganos, grupos de células sobre los que estudiamos sus características mediante un método de observación directa que no es estadístico. Pero entre esos conjuntos de seres vivos y los grupos sociales hay una diferencia esencial. Supongamos que yo mido la temperatura de un individuo. Bastará con que lo haga en una parte del cuerpo, en su interior. Si yo estudio un tejido, puedo limitarme, de nuevo, a observar una o varias células con el microscopio. En ambos ejemplos, al estudiar las funciones de un organismo podré conocer las funciones de todos los organismos de esa misma especie, puesto que dichos conjuntos están formados de elementos parecidos y casi idénticos. No sucede lo mismo en los grupos sociales y en los fenómenos, propiamente sociales, que tienen lugar en su interior. Las tendencias, las creencias y los pensamientos colectivos están representados de manera desigual y diversa en cada individuo, y en cada uno de ellos no encontraremos más que una parte o un aspecto concreto. Por esa razón sólo hay un modo de tener en cuenta a todo el conjunto. Consiste en agrupar todas las partes, contabilizándolas y procurando no olvidar ninguna de ellas, para volver a componer el conjunto. En el fondo, no hay más conjuntos reales que los grupos sociales que, de hecho, están formados por diferentes elementos. El resto de conjuntos solo son colecciones. Las especies vivas son colecciones de organismos, los órganos y los tejidos son colecciones de células, mientras que el organismo no es más que un individuo. Los grupos sociales son mucho más que eso y otra cosa diferente.

Partamos ahora de esta noción de grupos y de conjuntos, reales y consistentes. Aquí es donde comienzan las dificultades. Empecemos por estudiar a estos grupos; pero claro, no se trata de objetos sensibles con siluetas bien definidas. Por lo tanto, antes de estudiarlos tendremos que construirlos. Pero ¿cómo hacerlo?, ¿cómo construirlos si no los hemos estudiado aún? Se comprenderá mejor la naturaleza de un círculo vicioso como este, si recurrimos a algunos ejemplos.

La estadística oficial suele organizar sus observaciones dentro de límites espacio-temporales y, difícilmente puede hacerlo de otro modo. En ocasiones, dichos límites se corresponden con divisiones sociales reales como, por ejemplo, los nacimientos y los fallecimientos diferenciados por países o por años. Pero la población de un territorio no se corresponde con un grupo social, del mismo modo que un periodo de cinco o de diez años no es, normalmente, un periodo social definido. Es evidente que se necesitan puntos de referencia. Si miramos una parte del cielo a través de un telescopio o una muestra con un microscopio, comprendemos que una de las tareas del observador consiste en delimitar y dividir el objeto observado mediante un conjunto de líneas, separadas regularmente y que se cruzan entre ellas formando ángulos rectos. De esta manera, si el observador desplaza su instrumento de observación, de forma que esas líneas abarquen el conjunto o el detalle que sea de su interés, estas coincidirán más o menos con los límites reales del objeto que se propone estudiar. Pero el estadístico no puede modificar de manera artificial esos límites. Por ejemplo, cada cinco o diez años se elabora el censo de población y, el estadístico, considerando estos intervalos para las cifras de población, registra también los precios o los salarios. Pero puede que tal vez, en uno de los años intermedios se haya producido el inicio o el final de un movimiento que hasta entonces permanecía constante y que, por lo tanto, hará que se produzca una variación importante. Todavía las estadísticas oficiales no nos ofrecen más que aquello que pueden ofrecer y cabe destacar que se han hecho grandes progresos en ese sentido. Ahora conocemos, por ejemplo, los movimientos de población anuales, incluso trimestrales, y las oscilaciones de precios mes a mes. Pero pese a todo ello, ¡cuántos autores que trabajan con estas estadísticas continúan comparando datos tomados con 10, con 20 e incluso con 50 años de diferencia! ¿Qué clase de científico organiza sus datos con unos intervalos temporales así? Esta es la razón por la que Simiand insistía en la necesidad de estudiar un fenómeno en todo su desarrollo, de manera ininterrumpida, desde su comienzo hasta su final. Del mismo modo que un botánico no se limitará a observar el desarrollo de una planta el primer día de cada mes o todos los domingos, sino que lo hará día tras día, desde que germina hasta que florece.

En los anuarios estadísticos encontramos tablas que nos muestran cómo se distribuyen los matrimonios, partiendo de la diferencia de edad entre los cónyuges, pero en Francia nos limitamos a comparar mediante combinaciones de edad. Del siguiente modo: maridos de 20 a 25 años y mujeres de 20 a 25 años, maridos de 25 a 30 años y mujeres de 25 a 30 años. Sin embargo, la primera combinación incluye los casos en los que el marido tiene 20 años y la mujer también (diferencia de edad= 0), otros donde el marido tiene 25 años y la mujer tiene 20 (intervalo de edad= 5 años), y también, aquellos casos en los que el marido es mayor que la mujer o a la inversa, con una diferencia de edad de 1, 2, 3, 4 y 5 años. De este modo agrupamos en una misma categoría a matrimonios que, tanto desde un punto de vista biológico como social, representan uniones matrimoniales muy diferentes.

De la misma manera podríamos preguntarnos si los grupos de edad, tal y como los construyen los estadísticos, se corresponden con una realidad social. Trataremos de entenderlo a continuación. Cuando decimos que un animal tiene tal o cual edad, en principio dudamos si utilizar dicho término. A lo que hacemos referencia con ello, en todo caso, es que el animal ha llegado a una determinada etapa de su evolución biológica individual o, dicho de otro modo, de la evolución biológica natural de un representante de su especie. Pero el animal no sabe cuál es su edad, y si los hombres no fuesen otra cosa que animales, tampoco lo sabrían. Por otra parte, un individuo humano aislado, privado de cualquier contacto con sus semejantes y sin ninguna referencia acerca de la experiencia social, ni siquiera sabría que debe morir. Tal vez (si la memoria no se constituyese en la vida social) no recordaría que fue más joven de lo que lo es en el presente. Luego, ¿qué sentido tendría para él la noción de edad? Se trata, por lo tanto, de una noción social que se define por comparación con el resto de miembros del grupo y en todo caso, el grado de precisión con la que se nos impone se explica mediante la costumbre y las instituciones. Un estudiante turco me explicaba que en su país, donde el estado civil existe desde hace muy poco, la mayoría de hombres desconocen su edad y que, si se les pregunta, responden que no les interesa. Simplemente distinguen tres categorías: niños, adultos o ancianos y se reconocen en alguna de ellas.

Pero si nuestra edad está, en cierta medida, impuesta por la sociedad, no se sigue de ahí que los hombres o las mujeres de una determinada edad, expresada en un número de años, constituyan un grupo social definido. ¿Por qué razón? En primer lugar, porque en un grupo de esas características se permanece muy poco tiempo, puesto que cada año se aumenta un nivel y continuamente se pasa de una categoría a la siguiente. En segundo lugar, aunque se haya propuesto formar, por ejemplo, el conjunto de hombres de 40 años, si existen intereses y

preocupaciones que sean comunes a los jóvenes, los adultos o los ancianos, en realidad se trataría de grupos cuyos miembros tienen edades bastante diversas, y que se esparcen en intervalos de edad mucho más amplios. Por último, en función de la época, de las costumbres, de las instituciones y de la composición de la población, concedemos a la edad una mayor o menor importancia y, por lo tanto, la juventud, la edad adulta o la vejez merecen en cada momento opiniones y definiciones muy distintas. En otra época, un europeo de 50 años se pensaba joven para emprender un negocio en América, mientras que a esa misma edad en nuestro país, cerraba su comercio o se jubilaba. Pero tal vez esa situación ha cambiado después de la última guerra y quizá pueda cambiar de nuevo.

Así mismo, hay un rigor aritmético excesivo, un poco artificial y arbitrario, en las representaciones que nos ofrecen *las pirámides de población*. Estas están constituidas por filas de la misma altura y de longitud variable y, cada una de ellas corresponde a una categoría de edad: de 0 a 5 años, de 5 a 10, etc. En ellas, en la mitad derecha se representa a las mujeres y en la izquierda a los hombres. Si hacemos el ejercicio de superponer las proyecciones de dos pirámides de población, como por ejemplo la de Alemania y la de Francia, se puede constatar que, como cabría esperar, la pirámide alemana tiene una base mucho más ancha, puesto que la proporción de niños es más alta en ese país. Por otra parte, a partir de 25 años la pirámide de población alemana se encoge y, en ese mismo punto es sobrepasada por la proyección de la pirámide francesa, que pese a ser muy estrecha en la base, tiene sin embargo una proporción de personas mayores mucho más elevada que la alemana. En esa representación gráfica tenemos una traducción exacta de los datos numéricos, pero ¿cuál es su alcance desde un punto de vista sociológico? La cuestión estribaría en saber si los límites que separan la edad adulta de la juventud o la vejez de la edad adulta son los mismos, en la percepción común que se tienen de ellos, en ambos países. Y sobre esto podríamos dudar, ya que allí donde hay muchos hombres mayores, tal vez ellos se vean más jóvenes de lo que son a pesar de su edad. Pero si por el contrario hay muchos jóvenes, podemos suponer que un gran número de ellos ocuparán puestos y tendrán que enfrentarse a situaciones que, en otros lugares corresponderían a los adultos, por lo que tal vez ellos se vean a sí mismos, y el resto también los consideren, más mayores de lo que son en realidad, independientemente de su edad. Por otra parte, si tenemos en cuenta que estos dos países, uno está más al norte y el otro más al sur y que tienen una composición étnica diferente, puede ser que los hombres sean más precoces en uno, por ejemplo en Francia, que en el otro. Entonces, puesto que se alcanzaría antes la edad adulta, también se entraría antes en la categoría de ancianos, por lo que la población francesa sería aún más mayor y la población alemana aún

más joven de lo que pudiera parecer al observar la representación gráfica. Por último, ¿cómo puede obviarse la diversidad de clases sociales, de profesiones, de medios urbanos y rurales?, ¿es acaso igual la pirámide de población, en un mismo país, en la ciudad que en el campo, en la industria, en el comercio, en las profesiones liberales, entre las clases acomodadas o entre las clases pobres? Tengamos en cuenta, por ejemplo, que en los Estados Unidos la proporción de adultos es casi tan alta como en Francia, pero no porque la natalidad sea desde hace mucho tiempo tan baja, sino como consecuencia de la llegada de inmigrantes. Se trata de estar atentos a toda esa diversidad de condiciones y debería ser así como la investigación estadística seleccionase a sus grupos de estudio. Por todo ello, las pirámides de población nos ofrecen una idea tan esquemática y tan pobre como la que obtenemos de las pirámides de Egipto sobre el destino de los miles de hombres que las construyeron.

Lo mismo diría de otro procedimiento, en ciertos aspectos ingenioso y útil, que los estadísticos nos han propuesto recientemente con el fin de comparar la mortalidad, su media, de dos países. Cuando se observó que la mortalidad (media) o la mortalidad probable, digamos la tasa de mortalidad, era más baja en Alemania que en Francia, pudo animarnos a pensar, de manera ilusoria, que la esperanza media de vida era más alta en aquel país que en este. Teniendo en cuenta que la composición social por edades es diferente para ambos países, parece natural imaginar que el coste, en términos de mortalidad para una población más envejecida, sea más alto que para el caso de una más joven. Aunque se quiso excluir esa posibilidad de error y proceder, como se haría en física, eliminando o manteniendo constante un valor para medir el efecto del otro, y por esa razón se hizo abstracción de la composición según la edad. Se habrían podido limitar a comparar la tasa de mortalidad entre ambos países a igual edad. Pero se quiso ir más allá, de modo que se hizo el cálculo de una única tasa de mortalidad para toda la población en cada país y se rectificó teniendo en cuenta la distribución desigual según las diferentes edades. ¿Cuál fue el procedimiento? Consideremos una población tipo que, teniendo en cuenta ese modelo de distribución, ocupará una posición que se aproxime a la media entre los otros dos. Pensemos por ejemplo en Suecia. A continuación, a las diferentes categorías de edad de esta población tipo le asignaremos las tasas de mortalidad que correspondan a cada una de ellas, por ejemplo, primero en Francia y después en Alemania. El siguiente paso consistirá en calcular la tasa de mortalidad global para cada uno de dichos países. Se obtiene entonces que, mientras que la *tasa de mortalidad simple*, calculada siguiendo los métodos antiguos, es más elevada en Francia que en Alemania, la *tasa de mortalidad rectificada* es en Francia menos elevada que en Alemania

Más, ¿cuál es el alcance preciso de este resultado? La muerte es un hecho biológico, pero también es un hecho social. La tasa de mortalidad puede variar, no solo para el conjunto de la población, sino también en cada categoría de edad, más concretamente, en función de la composición de la población según las diferentes edades (como resultado de ciertas condiciones sociales). En un país en el que hay pocos niños y muchos adultos, estos últimos están menos expuestos a la posible competencia de los jóvenes, debido a su escasez, y por lo tanto disfrutan de una vida más tranquila. Por otro lado, puesto que los niños son pocos, tal vez se le preste más atención a sus cuidados, a su desarrollo físico y a su salud. En caso de que se hubiese modificado la composición de la población francesa en función de la edad, nada demuestra que las tasas de mortalidad por edad no serían más altas. Por lo tanto, no es legítimo conservar las tasas de mortalidad por edad y, al mismo tiempo modificar la composición de la población. De este modo hemos querido aislar el aspecto biológico de la muerte, atendiendo únicamente a los factores que determinan la vitalidad natural de una población, pero incorporando el estado de desarrollo de la medicina y de la higiene en cada país. Aunque la muerte, prescindiendo del hecho de la composición de la población según la edad, sigue estando condicionada por otros muchos factores sociales. La muerte es resultado de la vida, y la vida está condicionada por las condiciones económicas, por la organización familiar, por las instituciones y por las prácticas culturales. Todo ello habría que eliminarlo o, al menos, suponerlo constante para todos los países. Pero, ¿sería posible? Este problema nos plantea un buen número de paradojas. ¿Cuánto tiempo vivirían los franceses si, permaneciendo franceses, viviesen en las mismas condiciones físicas y sociales que los suecos? ¿Cuánto tiempo vivirían los alemanes si, permaneciendo alemanes, viviesen en las mismas condiciones que los franceses? Esto lleva a preguntarse, como observó Simiand a propósito de una comparación económica hecha recientemente sobre los niveles de vida en diferentes países, cómo viviría un camello si, permaneciendo camello, fuera transportado a las regiones polares y cómo viviría un reno si, permaneciendo reno, fuera transportado al Sáhara. Dicho de otro modo, todo sucede como si, para estudiar las características demográficas de un país, hubiera que partir de la población que no es la de ningún país, como si tratásemos con hombres que no nacen, no se casan o no mueren en una determinada región que, de algún modo, se define en función de sus tradiciones familiares, religiosas, jurídicas, económicas. Sin embargo, del mismo modo que el *homo economicus*, un tal *homo demographicus* es una abstracción demasiado cuidadosamente separada de la realidad, para que pueda enseñarnos algo sobre lo real.

Por descontado, no discutimos que estos procedimientos esquemáticos de representación, de análisis y de simplificación no sean útiles, ya que nos

proporcionan un conocimiento más claro acerca de nuestros datos y nos hacen descubrir en ellos una serie de relaciones a las que no teníamos acceso antes de haberlos elaborado mediante tales procedimientos. Lo relevante es saber, si lo más conveniente es iluminar los hechos de ese modo, en un número bastante elevado y limitado o si, por el contrario, sería mejor seleccionar otros hechos distintos. Aunque esto último no hace que debemos abandonar dichos procedimientos, no debemos tampoco limitarnos a ellos ni imaginar que hemos explicado la realidad, cuando la hemos substituido por una fórmula o por una representación que se ajusta un poco a ella. Porque entonces corremos el riesgo de superponer, a los grupos reales, unos grupos ficticios que, en principio, parecen corresponderse con los primeros pero solo porque, efectivamente, se trata de ellos mismos aunque privados de una buena parte de su contenido. ¿Estamos seguros, entonces, de que lo que se rechazó por motivos de simplificación no era lo esencial, aquello sin lo cual no podemos comprender la realidad?

Dijimos que la estadística era la ciencia de las medias y de las curvas, pero tal vez esta sea una definición insuficiente. En todo caso, es cierto que, por ejemplo, para eliminar las variaciones estacionales del desempleo hay que calcular las medias anuales y, para realzarlas, las medias mensuales. Simiand vio incluso en la sociología cuantitativa el equivalente del método experimental. Aunque aquí hay que adoptar precauciones adicionales, que no son necesarias en física, puesto que nosotros no disponemos de los mecanismos de control al alcance del físico, como la verificación material. Si el físico ha despreciado factores esenciales, el fenómeno no se produce. En nuestro caso, las cifras se dejan siempre combinar con otras cifras. Para estar seguro de que sus medias o sus índices se corresponden con la realidad, algo que en nuestro caso es más complejo que para la física, es necesario que el estadístico calcule también valores típicos complejos, como por ejemplo las medianas, con los cuartiles y los deciles, y que multiplique los tipos de cálculos y los cruces de datos. En lo referente a las curvas, estas deben reflejar todas las fluctuaciones del fenómeno y representarlo en todas sus fases, además de abarcarlo en toda su amplitud y en cada una de sus partes. Así será como representemos las variaciones salariales, mediante múltiples curvas, continuas cuando sea posible o yuxtapuestas, y tan numerosas como datos tengamos sobre diferentes grupos, en agricultura, industria, también sobre diversos tipos y formas de industria, grandes ciudades, ciudades de tamaño medio, ciudades pequeñas. Todo ello exige un esfuerzo de atención múltiple, a un mismo tiempo abstracto y concreto. Pero este método empírico es el único que permite mantener un contacto, lo más estrecho posible, con la realidad.

La opinión de muchos estadísticos con formación matemática es muy distinta. Está bastante bien expresada en estas dos frases que tomo prestadas de un libro, destacable en muchos aspectos, de Robert Gibrat, *Les inégalités économiques*: "Nuestra ley es esencialmente estadística. Simplemente, reduce las curvas de distribución económica a otra famosa curva, en forma de campana, también llamada de Gauss o de errores". Agrupar los hechos económicos y sociales, sus movimientos y sus variaciones, y reducirlos a una curva con la que los matemáticos están familiarizados, ¿ese es realmente el ideal de la investigación estadística? Ciertamente, el ingenio de los algebristas y los geómetras es grande. Disponen de numerosos procedimientos de ajuste y saben cómo introducir en sus fórmulas diferentes parámetros, de manera que, haciendo las modificaciones necesarias, son capaces de transformar las curvas observadas en curvas teóricas con formas conocidas. Así fue como el estadístico inglés Udny Yule mostraba, hace algunos años, que las curvas de población observadas durante más de un siglo, en tres países tan diferentes como Inglaterra, Francia y Estados Unidos, podían situarse, mediante ciertos ajustes aplicados sobre diferentes secciones o solapándose parcialmente, en una misma curva teórica con una forma perfectamente regular y que podía expresarse mediante una fórmula única. Aunque las curvas de ese tipo no nos enseñan nada acerca de los mecanismos internos en las conexiones de las diversas series de hechos, de donde surgen las variaciones sobre las que tales curvas nos ofrecen una representación aproximada. Sin embargo, será a partir del estudio de ese conjunto de conexiones cuando comenzará, y solamente a partir de ese momento, la investigación positiva. Podemos incluso temer, que las simplificaciones que se han aplicado para la fabricación de dichas curvas, hayan eliminado algunas de las características esenciales que necesitaríamos para estar sobre el buen camino de la explicación.

He aquí, por ejemplo, el estudio sobre la variación de precios, del que se han ocupado tan cuidadosamente muchos países desde hace algunas décadas. Nos dimos cuenta muy pronto que en los precios se manifestaban variaciones cíclicas de corta duración, de bajada, de estancamiento y después de subida, separadas por crisis periódicas con intervalos de 5, 6 o 7 años. Se producía un movimiento regular que parecía poder representarse por una curva como la que expresa las oscilaciones periódicas de un péndulo, esto es, mediante una curva sinusoidal. Nos concentramos en esta figura bastante simple, tal vez porque evocaba la idea de movimientos alternos y regulares, por debajo y por encima de un nivel medio, de una posición de equilibrio. Esta representación físico-matemática se impuso cada vez más, especialmente porque permitía agrupar en un modelo simple y uniforme toda la serie sucesiva de variaciones. Sin duda estaba claro que a través de dichas oscilaciones, y también entre ellas, se producía un cambio de

distinta naturaleza y de una duración más prolongada, en la subida o en la bajada, que bien acentuaba o bien atenuaba las oscilaciones, y que introducía otras irregularidades. Pero, en lugar de tratar de estudiar dicho cambio en sí mismo, se intentó eliminarlo, buscando transformar la pendiente de una montaña en un llano, en una línea horizontal. Se disoció, mediante el pensamiento, la pendiente del resto de accidentes (en este caso, las subidas y las bajadas, de su relación con la pendiente), y, antes que nada, para encontrar la pendiente, se partió de dichas subidas y bajadas.

Se determinó mediante el método de los mínimos cuadrados, o bien de otras maneras, y a partir de ahí, nunca ha vuelto a ser cuestionado. Se entendió que se trataba de una variación secundaria, sin relación con las oscilaciones y sin efecto sobre ellas. No cabe duda que, en el *Cours de statistique* de Albert Aftalion se nos dice: "Siguiendo algunas teorías bastante consolidadas, el movimiento a largo plazo de los precios, se produciría mediante grandes ritmos alternos". Pero añade el autor: "Podemos dudar de la existencia de esos grandes ciclos. Hasta el momento, solo se han identificado dos y medio (desde finales del siglo XVIII), lo cual es muy poco para poder establecer una regularidad duradera". Para Simiand, por el contrario, una o dos repeticiones serán suficientes para confirmar la existencia de esos grandes ciclos. Eso fue lo que él pudo observar mediante su complejo procedimiento. Dos experiencias podían resultar suficientes, siempre que se integrasen en un conjunto denso, compuesto por numerosos factores, medidos de manera precisa. Por su parte, Albert Aftalion se concentra simplemente en las subidas y bajadas alternas, en la curva simplificada que, difícilmente, podrá ser característica, puesto que no representa más que uno o dos casos. De esta manera, toda la atención se centra en los movimientos cíclicos del corto plazo, y con ellos, nace una vasta literatura al respecto. Se hacen esfuerzos para producir teorías, apartando dichos movimientos del movimiento más amplio del que se extraen y se trata de explicarlos de manera aislada.

Simiand, por el contrario, mediante una observación directa que abarcaba más de un siglo, y sin partir de variaciones cíclicas cortas, ha determinado la existencia de movimientos periódicos de larga duración, para el caso de los precios. Él pudo comprender las características de tales movimientos, partiendo de la condición de situarlos en el conjunto de la evolución económica, tomada en toda su dimensión, es decir, introduciendo de nuevo en su objeto de estudio todo aquello que previamente se había descartado. Desconozco todavía cómo explicaba él las variaciones cíclicas en el corto plazo, a lo que dedicó su último curso del Colegio de Francia. Pero, salvo que yo me equivoque gravemente, creo

que fue, considerando los movimientos en el largo plazo y remplazándolos, como comprendió su funcionamiento.

Lo que yo sé, en todo caso, es que para él, las variaciones cíclicas no se reducían a oscilaciones, por debajo y por encima de una posición de equilibrio, sino que vio las fases sucesivas de una evolución orgánica, una especie de avance, de tal modo que cada ciclo o cada fragmento de un ciclo no se correspondía con la reproducción automática en un mismo sentido (en el caso de un ciclo completo), o en el sentido opuesto (si se trata de la mitad de un ciclo) del movimiento anterior. No había una vuelta periódica a un mismo equilibrio, sino el paso de un desequilibrio a otro desequilibrio o, si se quiere, de un equilibrio inestable a otro equilibrio inestable diferente del primero, por la única razón que el primero se produjo anteriormente y que, en cierto modo, lo había preparado, siendo ésta su condición. Sobre una evolución de este tipo, ninguna representación esquemática podría ayudarnos a hacernos una idea precisa.

De hecho, la utilización de los métodos estadísticos en sociología cuantitativa, debe pensarse de un modo diferente. Los métodos no nos proporcionan teorías, sino instrumentos de observación y de comparación que son a la vez, precisos y objetivos, y es en ese sentido en el que deberán desarrollarse. Desde esa óptica, y sin adentrarnos en los detalles técnicos, deben examinarse los procedimientos tan frecuentemente recomendados y aplicados, desde hace algún tiempo, bajo el nombre de índice y coeficiente de dependencia o coeficiente de correlación. Su objetivo es el de sustituir la comparación visual de curvas empíricas por el cálculo de un número único, comprendido, por ejemplo, entre -1 y +1, donde +1 representa el grado máximo de correspondencia entre dos series de datos, 0 representa la ausencia de cualquier relación, y -1 el grado máximo de oposición (lo que implica también la existencia de un grado de dependencia). La ventaja que presenta un método de este tipo, es que substituye una apreciación subjetiva y cualitativa (que puede variar entre dos observadores), por un resultado objetivo y preciso. De una manera automática, el cálculo responderá a las preguntas: ¿existe dependencia entre estas dos series?, ¿cuál es el grado de dicha dependencia? No será necesario entonces observar las curvas durante mucho tiempo, ni tampoco demostrar ingenio, intuición o delicadeza, puesto que cualquiera que opere con dichos cálculos podrá llevar a cabo una suerte de selección de aquellos hechos que se ponen en relación y el resto. En suma, un trabajo de laboratorio, de equipo, cuya rentabilidad está asegurada.

Sin embargo, en primer lugar, debe asumirse que, sobre el valor mismo de los diferentes coeficientes, los matemáticos no están de acuerdo, o más bien, parecen haberse puesto de acuerdo en reconocer que no han hallado todavía un

coeficiente de correlación que resulte completamente válido. Sobre el coeficiente de correlación de Bravais-Galton, tan frecuentemente utilizado, Maurice Fréchet, a partir de una investigación llevada a cabo junto a algunos de los estadísticos más reconocidos por sus competencias matemáticas, ha presentado recientemente, en el último Congreso del Instituto internacional de estadística, la siguiente proposición, en la que constata que: "el ámbito de validez del uso del coeficiente de correlación es notablemente más limitado de lo que en principio se había asumido; que ningún valor del coeficiente de correlación, ya fuera +1 o -1, garantizaría por sí mismo la causalidad, puesto que ésta debe probarse mediante razones de un orden diferente".

Aunque lo que nos parece especialmente discutible es el principio mismo sobre el que se apoya este método. Se trata de destacar lo que, desde Stuart Mill, se llama variaciones concomitantes en dos series. No se limitan a comparar las "primeras diferencias", o las "variaciones adyacentes", es decir, las variaciones en los elementos que se encuentran unos junto a otros, sino que es posible considerar la variación de cada serie, en relación a su media, en cada una de ellas. Por otro lado, se elimina la diferencia de intensidad en las variaciones de dos series, dividiendo las desviaciones por la desviación cuadrática media de la serie (lo que los anglosajones llaman la *desviación estándar*). La atención se sitúa, por tanto, en el sentido general y la importancia relativa de las variaciones, y todo ello está bien. Pero, de entrada, tal y como se ha señalado a menudo, cuando dos series varían de forma concomitante, efectivamente existe covariación, aunque no por ello puede decirse que exista correlación. Tal vez haya una tercera serie, en relación y separada de las dos anteriores, que sea la que impulse los movimientos de ambas. En alguna ocasión encontré que existía una covariación bastante fuerte entre el número de suicidios y los datos sobre el comercio, relativos a las exportaciones francesas, entre 1870 y 1919. Tal vez se deba a que ambos hechos tienen relación con el progreso industrial.

Por otra parte, y esto es fundamental, el inconveniente principal de un método como este, es que ignora el orden en el que las variaciones se suceden a lo largo del tiempo, y esto no es desdeñable. Podríamos desplazar conjuntamente dos variaciones concomitantes de este tipo, llevarlas del centro al inicio o al final, podríamos incluso, considerar el orden de sucesión inverso y empezar por el final, y aun así, el coeficiente de covariación seguiría siendo el mismo. A pesar de que la secuencia de los hechos sería distinta, ya que nos hallamos en un ámbito en el cual las sucesiones de los fenómenos no son reversibles. Se ha constatado, por ejemplo, que cuando los precios aumentan, los salarios aumentan, pero cuando los precios bajan, los salarios no lo hacen. Por esa razón, si nos desentendemos del orden de sucesión ¿cómo distinguiremos las series de

variaciones que son la causa de las que son el efecto? Aunque podemos, sin duda, hacer la comparación entre variaciones a intervalos de uno o dos años, y esa distancia nos indicará entonces qué serie de variaciones es la anterior, y por lo tanto la causa. Así fue como lo hicieron los estadísticos de Harvard con sus tres curvas, de movimientos bursátiles, de cifra de negocios y de movimientos monetarios. Pero sabemos que estos métodos predictivos no han tenido apenas éxito cuando se han aplicado al pasado. Por otra parte, entre las series se producen acciones y reacciones recíprocas, de manera que tal variación, siendo causa en un momento, será efecto en otro, con respecto a las variaciones de la serie vecina. De ello, solo podemos estar prevenidos mediante una observación directa y continua, que siga el orden mismo por el cual los fenómenos se suceden de un año al siguiente. Esto quiere decir, que un método que haga caso omiso de dicho orden, puede que tal vez esté dejando escapar lo esencial.

No cabe duda de que el objeto fundamental del método estadístico es buscar las concomitancias entre series de variaciones. Pero debe entenderse bien que una variación puede tener un significado muy diferente, incluso si mantiene el mismo sentido e intensidad, en función de la evolución de la fase en la que se produzca. En un periodo de crecimiento económico puede producirse un aumento de bienes manufacturados, ya que las oportunidades comerciales aumentan y los precios son ventajosos para los fabricantes. Pero esto también podría suceder, y de hecho sucede, tal y como lo mostró Simiand, durante periodos de crisis, precisamente porque los precios disminuyen y se trata de recuperar, mediante el volumen de bienes producidos, aquello que se pierde en términos de precio unitario. En esos casos, la covariación no nos proporciona ninguna información y necesita ser interpretada. Un avance en ese sentido sería disponer, al mismo tiempo, de una comparación entre dos series, con la ayuda de varios índices o coeficientes, como por ejemplo, un índice de dependencia que se centrara en el detalle de las variaciones adyacentes o un índice de correlación que se ocupara del conjunto; disponer también, de comparaciones entre más de dos series observadas al mismo tiempo, mediante las cuales podríamos observar las variaciones, en función de su orden y su intensidad, tanto de su movimiento como de su ritmo. Ciertamente, en este sentido las matemáticas nos ofrecen una enorme cantidad de procedimientos distintos, que pueden adaptarse a los casos más variados y complicados. Pero, habida cuenta de la maraña que compone los hechos sociales, ¿no corremos el riesgo de perdernos en fórmulas abstractas?, ¿no es más conveniente fijar la atención en las curvas empíricas, que ya son lo suficientemente numerosas y variadas, y que demandan todo el esfuerzo posible por parte del observador? El instrumento matemático es potente, pero tal vez obtengamos mejores resultados si determinamos, en primer lugar y mediante otros métodos, los vínculos que nos interesa esclarecer y los límites del campo

que exploramos. Sería una ilusión que tendría graves consecuencias, si creyésemos que vamos a descubrir esos vínculos, en un campo como el nuestro, mediante el cálculo, midiendo todo aquello que se nos presente, de todas las formas posibles y en todos sus aspectos. Puesto que la cantidad de combinaciones posibles es ilimitada, no las agotaríamos jamás y nunca alcanzaríamos nuestro objetivo.

Si recordamos ahora que el método estadístico en sociología debe aplicarse a conjuntos o grupos reales con el fin de revelar sus características, que solamente existen para dichos conjuntos y que, a menudo, no se encuentran en ninguno de los individuos o incluso en ninguno de los grupos parciales de individuos que podamos elaborar, parecerá que, entonces, el peligro está en tratar de reconstruir el conjunto a partir de alguna de sus partes. Así será, por ejemplo, retomando el caso de la demografía, la interpolación que, a partir de una parte elegida arbitrariamente (con frecuencia a partir de los últimos 10 o 20 años o, en ocasiones, a partir de movimientos de población producidos en los dos o tres años anteriores) determina cuál será la población de Francia, o de otro país, de un continente o del mundo, en el año 1950 o en el año 2000. Pero más bien, solamente partiendo de la población observada durante una cincuentena de años, podremos conocer exactamente el carácter de las tendencias demográficas durante una pequeña parte de ese largo periodo. Así de nuevo, partiendo de las tasas de natalidad y de mortalidad de las mujeres solteras y en edad fértil, en nuestra población actualmente, se concluye que, pasadas unas décadas, en lugar de 1000 mujeres solteras en edad de contraer matrimonio, nada más habrá 900, 800, 600, como si la discusión no fuera, en realidad, si en las próximas décadas, esas tasas de natalidad y de mortalidad no cambiarán. Se hacen enormes esfuerzos de mucho ingenio en elaboraciones similares, cuando, en realidad, no aumentan mucho nuestro conocimiento y solo sirven para constatar, tal vez de un modo sorprendente, aquello que ya sabíamos y que solo es aplicable al corto periodo al que los datos hacen referencia. Aprenderíamos mucho más si nos concentrásemos, sobre todo, en explorar los periodos más antiguos, si ampliásemos las observaciones en el espacio, si descompusiésemos los conjuntos globales y confusos que representan las poblaciones, en grupos más restringidos, que no arbitrarios, tratando de encontrar, mediante ensayos y cruces, las articulaciones naturales de las sociedades. Leo a menudo, con mucho respeto hacia los científicos que los firman, artículos de estadística matemática que se publican en revistas como *Metron* y *Econometrica*, y debo admitir que estoy bastante sorprendido de lo poco que suelen encontrarse, en ellos, realidades positivas. También me sorprende la indiferencia que parecen despertar habitualmente los hechos más simples, en su amplitud y su detalle, y en su extraordinaria variedad. Aunque es cierto que, entre los demógrafos y los

historiadores de la economía, falta muy a menudo ese mínimo de cultura matemática sin la cual, la acumulación de datos, no es más que erudición un tanto vana. La estadística matemática nos sería mucho más útil y progresaría más deprisa en sus aplicaciones en las ciencias sociales, si aquellos que se dedican a ella prestasen más atención a los resultados contemporáneos de la ciencia positiva.

Decía Renan, a propósito de la lógica y de las matemáticas: "No se trata de ciencias a parte y progresivas, sino de un conjunto de conceptos fijos. No nos enseñan nada, pero son capaces de analizar lo que ya sabemos. No negaremos que hay ciencias de lo eterno y de lo inmutable, pero situémoslas, claramente, fuera de cualquier realidad". Yo no haría mía una afirmación como esa, que no destaca lo suficiente que las matemáticas están, en cualquier caso, arraigadas en lo real. Las matemáticas pueden ser una herramienta con un potencial incomparable, siempre que se adapten a los avances de las ciencias positivas. Aunque muy a menudo sucede al contrario. Da la impresión que nos interesamos, preferentemente, por aquellos problemas para los que la estadística matemática nos ofrece fórmulas bien probadas y sencillas. Pero, qué diríamos, por ejemplo, de un médico que acude a visitar a un enfermo aquejado, tal vez, de problemas nerviosos después de mucho tiempo, puede que neurasténico o ciclotímico y que se limitase a utilizar los sofisticados instrumentos que lleva consigo, como un termómetro, aparatos que midan la presión o que sirvan para observar el interior del cuerpo. ¿No sería esencial que, al contrario, en primer lugar, se preguntase e intentase construir una visión de conjunto de la evolución de la enfermedad hasta ese momento? Porque, de hacerlo de ese modo, las observaciones precisas que pueda llevar a cabo con su instrumental, estarán mejor dotadas de sentido. En las ciencias sociales y también en materia social, será adaptando la estadística matemática a los problemas que nos plantea la investigación positiva, y reclamándole medios para ser más precisos en nuestras mediciones e instrumentos de expresión más exactos, como tendremos alguna posibilidad de avanzar. En lo referente a las explicaciones, hay razones para pensar que no debieran ser de tipo matemático, ya que nos encontramos en un campo donde, ciertamente, los hechos son más complejos que aquellos en los que se concentra la biología.