

IX Reunión ASEPELT-ESPAÑA
Santiago de Compostela,
23 de junio de 1995

La investigación estadística como aproximación a la realidad

José Aranda Aznar
Director Gral. Estadísticas de Población y Hogares. INE

Cuando, como ha ocurrido durante los dos últimos días en este foro científico, se presentan ponencias y comunicaciones basadas en datos estadísticos, quienes nos ocupamos de la producción de dichos datos sentimos, por un lado, la satisfacción de constatar la utilidad de nuestro trabajo pero, por otro lado, también sentimos la lógica preocupación de quien conoce las limitaciones de su tarea.

El que la investigación estadística puede, en teoría, aportar un conocimiento riguroso de la realidad es algo fuera de toda discusión. Analicemos el ejemplo de las estimaciones de resultados electorales. Aquellas que se basan en el recuento de papeletas realizado en una muestra de mesas electorales, proporcionan una estimación muy precisa del verdadero resultado, cuestión constatable por la sencilla razón de que el recuento oficial exhaustivo de los votos puede considerarse un «resultado verdadero», dadas las garantías democráticas que protegen el escrutinio de los votos. Este ejemplo de las estimaciones de resultados de una consulta electoral es uno de los pocos en los que se puede comparar el resultado de la estimación con un resultado verdadero, puesto que cuando se investigan otros fenómenos (rentas, consumo, ocupación, desempleo, equipamiento, etc.) no se dispone de un resultado verdadero que, con carácter incontrovertible, pueda indicarnos exactamente el sesgo en que han incurrido nuestras estimaciones.

Con objeto de señalar los problemas con que tropieza la investigación estadística, voy a continuar con el ejemplo de las estimaciones de resultados en una consulta electoral. Cuando la estimación de dichos resultados no se basa en un hecho cierto como el citado escrutinio de votos en las mesas electorales seleccionadas en la muestra, sino, por ejemplo, en las entrevistas realizadas en las puertas de los Colegios Electorales, la precisión de las estimaciones varía muy notablemente.

Las razones que llevan a sesgar los resultados de las investigaciones por entrevista obedecen, por un lado, a los propios errores en la selección de las mesas electorales y al modo de selección de las personas que salgan de ejercer su derecho al voto en los Colegios Electorales donde dichas mesas se ubican. Estos errores, medibles por procedimientos estadísticos, pueden reducirse al máximo probando la representatividad de las secciones mediante resultados de consultas electorales anteriores e implantando estrictos sistemas de selección sistemática de los votantes, que eviten la propensión por parte de los encuestadores a entrevistar a determinado tipo de personas.

Por otro lado, existen otras razones para que la información obtenida aparezca sesgada, fundamentalmente las negativas a declarar el voto emitido y el falseamiento de la respuesta relativa al partido al que efectivamente se ha votado. El primero de los problemas puede aliviarse también mediante un riguroso procedimiento de sustitución de quienes se han negado a declarar su voto. Sin embargo, el segundo de los problemas, relativo al falseamiento de la respuesta, tiene una solución difícil y sólo sobre las base de las experiencias anteriores resulta posible aplicar índices de corrección a los datos agregados.

Para que se pueda tener una idea de la envergadura de las correcciones aplicadas en las últimas Elecciones Municipales, gracias a la amabilidad y confianza de algunos directores de grandes Consultoras que me han facilitado los datos, puedo decir que el voto declarado a Herri Batasuna se redujo en más de un tercio; el voto al PSOE se aumentó en más de un 10 por ciento; el voto al PP se mantuvo directo excepto en algunas provincias del País Vasco y Cataluña, donde se aumentó en un 15 por ciento; y el voto a IU se disminuyó en más de un 15 por ciento.

Debido a estos índices de corrección, verdaderamente importantes, las estimaciones facilitadas a las 8 en punto de la noche del pasado 28 de mayo alcanzaron una precisión considerable, pero podríamos imaginar el desastre que hubiera supuesto computar como bueno el voto declarado por las personas entrevistadas, sin introducir los citados factores de corrección.

Cabría especular sobre las múltiples razones que pueden llevar a los ciudadanos entrevistados a falsear sus respuestas y no sería muy aventurado suponer el temor de muchos residentes en el País Vasco por confesar que han votado a un partido «españolista», la resistencia actual de algunos electores por confesar su voto al PSOE, etc., pero lo más destacable del procedimiento de corrección utilizado es que únicamente resulta posible por existir consultas electorales anteriores que precisamente han puesto de relieve este tipo de sesgos al permitir, y lo repito nuevamente, comparar las estimaciones directas con un resultado que puede calificarse de verdadero.

Aun así, procedimientos como el señalado no están exentos de peligro. Efectivamente, las condiciones por las cuales se falsean las respuestas pueden variar (desaparecer el temor a los violentos nacionalistas, variar el sentido del voto de castigo, etc) y, de esta forma, los factores de corrección conseguirían un efecto contrario al buscado. De hecho, en multitud de fenómenos investigados varían las condiciones por las que las personas entrevistadas pueden falsear sus respuestas. Mejoras en la inspección y gestión tributaria darán origen a declaraciones de rentas superiores; un cambio en la legislación en materia laboral tendrá repercusiones en las estadísticas administrativas sobre ocupación y desempleo; alteraciones en la consideración social de determinadas situaciones (divorcio, minusvalías, etc) serán causa de la ocultación o el afloramiento de dichas situaciones.

De este modo, muchas de las mejoras o empeoramientos que se pueden poner de relieve en la situación de determinados campos de preocupación social obedecerán, por una parte, a auténticas mejoras o empeoramientos pero, por otro lado, vendrán influidas por cambios como los mencionados que alteran la realidad de un modo no medible por carecer, a diferencia de lo que ocurre en el ya reiterado tema electoral, de auténticos valores con los que contrastar las estimaciones.

Como se sabe, la investigación estadística está en disposición de cuantificar los errores debidos al muestreo y puede aliviar el problema de sustituir personas ausentes o que se hayan negado a colaborar mediante la aplicación de rigurosos sistemas de sustitución, imputación o ponderación de datos. No obstante, este alivio puede resultar insatisfactorio en la medida en que algunas unidades seleccionadas en primera instancia, como por ejemplo familias con un elevadísimo nivel de renta o empresas líderes en el sector, no encuentren unidades equivalentes aunque se trate de familias que residen próximas a las titulares de la muestra o de empresas del mismo sector y parecido tamaño a las seleccionadas en primera instancia. Pero el problema que deja inerte a la investigación estadística es el del falseamiento de respuestas.

La consagración del secreto estadístico en las Leyes de Estadística, que obliga a la publicación de datos agregados que impidan la identificación de informaciones individuales, no ha resuelto el problema del falseamiento de la realidad. Se está lejos de conseguir esa especie de acuerdo social mediante el cual los ciudadanos, por un lado, mantienen cuantas conductas irregulares les permitan sus propias conciencias o la permisividad del Estado pero, por otro lado, conscientes de la importancia que tiene un conocimiento riguroso de la sociedad, confiesan dichas conductas sin temor alguno, confiando plenamente en el secreto que protege sus datos individuales.

Los ciudadanos, quizás por no estar demasiado acostumbrados a que la información se utilice en su beneficio, saben que, si bien los datos individuales que

han facilitado no se pueden utilizar directamente en su contra, sí que resulta posible tener en cuenta datos agregados del colectivo al que ellos pertenecen y que sean tomadas medidas que afecten negativamente a todos los componentes de ese colectivo y, por tanto, a ellos mismos.

No obstante, y a pesar del evidente falseamiento que se da en las respuestas, el análisis de datos permite acercarse al conocimiento de esa realidad que tiende a ocultarse. En algunos casos, como ocurre con encuestas continuas como EPA y EPF, se considera que, con independencia de la credibilidad de los valores totales, la evolución que ofrecen de los fenómenos investigados (ocupación, desempleo, consumo, rentas, etc.) es correcta siempre que se considere, lo que no parece muy aventurado, que los recelos de las personas entrevistadas se mantienen estables en el corto y medio plazo. Para aproximarse a esos totales, el análisis de datos profundiza en los valores arrojados por distintas fuentes e intenta integrar la información, como ocurre con la labor de síntesis que realiza la Contabilidad Nacional donde, como se sabe, hay valores de producción de distintos sectores que es preciso aumentar en más de un 20 por ciento para hacerlos coherentes con las correspondientes cifras de consumo e incluso cifras de empleo que es preciso variar para ajustarlas, vía productividad, a las correspondientes cifras de producción. Este es el único camino posible para aproximarse a la realidad, dado que una Oficina de Estadística nunca puede realizar hipótesis, como la mencionada para el caso de las entrevistas electorales, para corregir los datos declarados por los ciudadanos entrevistados, puesto que ello, además de ser siempre arbitrario por carecer de valores verdaderos a los que referirse, abriría un peligrosísimo camino hacia posibles manipulaciones de la información. Los profesionales de la Estadística han de limitarse a hacer lo mejor posible su trabajo, reduciendo al máximo las incidencias, y deben dejar constancia de cuantas sustituciones e imputaciones de datos se hayan visto obligados a realizar.

Para terminar, quiero hacer una breve reflexión sobre el tratamiento dado a las ocultaciones de información hace dos siglos, en una época incipiente de la investigación estadística pero marcada por el signo de la ilustración que caracterizó el final del siglo XVIII.

Resulta muy interesante la lectura de la Advertencia del Censo de Floridablanca de 1787. El redactor de dicha Advertencia duda sobre el aumento de población entre 1768 y 1787, estimando que debería ser un 45 por ciento superior. Seguidamente da como razón de esta subestimación de población *«el cuidado con que los pueblos y sus vecinos procuran disminuir el número de sus habitantes, temerosos de que tales numeraciones se dirijan a aumentar las cargas de los servicios personales o de los tributos»*.

Quiero citar también el comentario que acompaña a la publicación de resultados del Censo de Frutos y Manufacturas de 1799, que pretendía nada menos que dar *«una razón de los frutos y de las manufacturas que ha producido cada Provincia en dicho año; sus precios corrientes, la cantidad que ha consumido y sobrado; la proporción que hay entre los productos y la población, y entre esta y la extensión territorial; y asimismo los lugares donde se hallan establecidos los artículos principales de la industria»*.

Tan ambicioso proyecto no debió ser un éxito por la Advertencia que figura en la citada publicación: *«La poca exactitud que se encuentra en muchos de los estados remitidos por los Intendentes; las faltas que se han notado en algunos, y la obscuridad que han presentado otros, hicieron demasíadamente trabajosa la redacción de este Censo, el cual carece para ello de la certeza que desearán los que le leyeran»*.

De estos comentarios, similares a otros muchos que aparecen en publicaciones históricas, donde se deja constancia del problema de la deformación de la realidad, me impresiona un detalle que a personas con sensibilidad distinta ante este problema podría pasarles desapercibido y que no es otro que la *naturalidad* con la que se reconoce la tendencia de la población a ocultar determinado tipo de situaciones e incluso a ocultar la propia existencia de personas, como en el caso de los Censos de Población.

Dentro de las actuales pautas culturales, con la agresividad imperante entre fuerzas políticas y sociales, con el talante que caracteriza a los medios de comunicación, no sería difícil adivinar las reacciones que se producirían ante comentarios semejantes a los referidos, aplicables ahora a poblaciones que tienden a sobredimensionarse para aprovechar los beneficios que se derivan de mayores cifras de población, a empleados sumergidos que cobran fraudulentamente subsidios de desempleo, a profesionales que ocupan ficticiamente a familiares para aumentar gastos deducibles o a empresarios que ocultan la totalidad o parte de su actividad.

El reconocimiento abierto de este tipo de circunstancias, por otra parte conocidas por todos, generaría sin duda reacciones extraordinarias y los responsables de las Oficinas de Estadística solemos limitarnos a sugerir la existencia de estos problemas cuando insistimos en que se colabore de un modo veraz y cuando reiteramos hasta la saciedad la existencia del secreto estadístico que protege los datos individuales que nos son confiados.

El problema a que nos hemos de enfrentar en ese trabajo nuestro de permanente aproximación a la realidad es netamente cultural, de confianza de los ciudadanos en sus administradores, de espíritu solidario que no lleve a los distintos grupos sociales

a intentar conseguir ventajas sobre los demás. Estamos desgraciadamente lejos de esa situación y se da la paradoja de que, por un lado, se nos oculte información y, por otro, se nos exija la mayor precisión en nuestras estimaciones.

En este sentido, y como una llamada a la racionalidad y a la tolerancia, voy a leer lo que se decía en el citado Censo de Frutos y Manufacturas. *A pesar de los defectos indicados, este Censo ofrece las mayores ventajas al Gobierno, por cuanto las luces que se deducen de él, evitarán los errores en que se puede incurrir por su falta. Y poco después añade, es de desear que las personas ilustradas, y que tengan conocimientos capaces de rectificar los que se contienen en este Censo, en lugar de emplearlos en dar materia a la censura, los manifiesten con sinceridad y buena fe, seguros de que en ello harán un servicio importante al Estado, de que S.M. los recibirá con la gratitud propia de su bondad y de que sus nombres no caerán en el olvido.*