Metodología de Encuestas
Vol 4, Núm 1, 2002, 75-77
ISSN: 1575-7803

¿QUÉ SON LOS MODELOS DE SUPERPOBLACIÓN?

Presentación

Imaginemos, por un momento, que estamos en pleno proceso de diseño para la realización de una encuesta sobre estimación de voto. El objetivo es estimar los resultados de las siguientes elecciones según determinados criterios como, por ejemplo, porcentaje de votos que recibirá el partido $\tilde{N}$ en la región Q. Para ello, en esta fase de diseño, se plantea la construcción de un cuestionario y la selección de una muestra.

La intención de voto cuenta con cierta varianza en la población, es decir, no todos los habitantes tienen la misma intención. En caso contrario, los resultados mostrarían una coincidencia total: toda la población vota al partido *G*, con lo que, incluso, desaparece la necesidad de realizar una encuesta. Debido a la existencia de varianza, es necesario concentrarse en resolver el problema de a quién entrevistar. El método de extracción de la muestra no es un tema inocente; si no se hace bien, los resultados de la muestra pueden estar apuntando en una dirección que no es la misma de la población. El muestreo, entonces, aborda el problema de la representatividad de los resultados muestrales, a partir de las operaciones que se realizan en esa porción de la población. Se desea reducir lo que ha venido a llamarse *error de muestreo*, de tal forma que la encuesta permita dar en la diana de los resultados electorales.

Supongamos resuelto el problema de la muestra. Tenemos ya a las personas seleccionadas, son quienes van a manifestar su intención de voto para la encuesta. Otro grave problema se refiere a las características de esa herramienta de medida que llamamos cuestionario. Es posible que no estemos preguntando lo que deberíamos preguntar. Tal vez sospechemos que las personas seleccionadas no querrán confesar su intención de voto y ello nos obliga a poner en marcha estrategias más o menos sofisticadas para encontrar el verdadero valor de lo que nos interesa ¿a quién piensa votar?. En cualquier caso, el cuestionario se enfrenta al imponente problema de lo que ha venido a llamarse

error de medida ¿Realmente estamos midiendo aquello que hemos planteado en los objetivos de la encuesta?.

Pero aún hay más. Las personas son entrevistadas a lo largo de una semana. La última recibirá la visita o la llamada de teléfono ocho días antes de los comicios. ¿Su intención de voto no sufrirá cambios?. El día siete antes de las elecciones, una entrevista televisada puede ser decisiva para variar la intención de mil doscientos electores. La opinión es un ser cambiante. Si interrogáramos por segunda vez a la misma persona seleccionada en la muestra, utilizando el mismo instrumento de medida, aún siguiendo un método de muestreo y un cuestionario que rayan la perfección... ¿Obtendríamos el mismo resultado? En otros términos, los valores que suministran las personas cuando son entrevistadas en las encuestas ¿Son entidades fijas? ¿No dependen de ningún factor variable? ¿No podría ser que se comportaran más como variables que como constantes?.

La forma tradicional de abordar los escollos que encuentra en su camino un equipo de investigación por encuestas, se ha venido centrando en solucionar, sobre todo, el error de muestreo y, en segundo lugar, el error de medida. Pero qué hacer con este último, algo que podríamos catalogar como *error de representación del dato*. La respuesta de una persona a un ítem en un cuestionario representa a un conjunto amplio: el abanico de posibles respuestas que esta misma persona podría realmente suministrar ante la misma cuestión. ¿Pero cómo abordar esta nueva perspectiva? ¿Cómo hacer las estimaciones considerando que las personas no suministran constantes sino variables?.

El error de muestreo, la distancia entre lo que se ha encontrado en la muestra y el valor verdadero en la población, es lo que suele acaparar la atención en las investigaciones por muestreo. Más concretamente, lo que ocupa es la varianza de ese error (la variación que se encontraría en el conjunto total de las muestras que podrían ser extraídas en las mismas condiciones), conocido como error cuadrático medio. Cada vez más, el error de medida, viejo acompañante de los psicómetras, se abre paso y va ocupando mayor protagonismo. Toca ahora llamar la atención e insistir en la circunstancia de que tal vez no tenga mucho sentido hablar del "valor verdadero" en la población, puesto que ésta es un ser cambiante que suministrará valores en un intervalo, tal vez ajustables a un modelo aleatorio. De esta manera, la población fija objeto del estudio representa, a su vez, a una población más extensa, variable, que recibirá los resultados de la estimación.

La solución no es nueva, el sistema conocido como "modelos de superpoblación" lleva ya muchos años operativo y, supuestamente, al alcance de quienes quieran considerarlos para las estimaciones a partir de muestreos en las encuestas. No obstante, la experiencia muestra que todavía se trata de un territorio muy poco conocido, tal vez debido al aumento de la complejidad matemática de su tratamiento, que no de su concepto.

En lo que sigue, adjuntamos dos textos escritos por dos autores que han merecido sobradamente nuestra confianza para realizar este encargo: Román Pérez Villalta, profesor de Economía Aplicada, en las licenciaturas de Economía y de Investigación de Mercados en la Universidad de Sevilla; y Gonzalo Sánchez-Crespo, delegado del Instituto Nacional de Estadística de Santander. Ambos se han enfrentado a los modelos de superpoblación como investigadores, como autores o como docentes, como creadores de herramientas o como generadores de textos. Sea como fuere, su valiosa experiencia en este campo nos ha llevado a pedirles la tarea de escribir para Metodología de Encuestas unos textos que no nos presentan nada nuevo, pero que, esperamos, permitan a un sector im-

portante de los profesionales e interesados en la investigación por encuestas, adentrarse en la dimensión de las superpoblaciones o, como dice Sánchez-Crespo, de las *supra*poblaciones.

Para obtener el máximo rendimiento de este breve trabajo de presentación, sería deseable que el lector contara ya con cierto conocimiento básico sobre nociones de muestreo en poblaciones finitas y sobre notaciones matemáticas. Los textos mencionados por Pérez Villalta, de Azorín y Sánchez Crespo, de Fernández y Mayor o de Mirás, son ejemplos muy recomendables.