

Encuestas y muestras: tres dudas y tres reglas

Hace unos años la empresa Gallup realizó en Estados Unidos un experimento muy curioso: una encuesta sobre las encuestas. Resultó que la mayoría de los entrevistados creía que las encuestas servían para saber lo opinión de la gente sobre asuntos de actualidad, o para predecir los resultados de las elecciones. Pero también, contradictoriamente, la mayoría de los encuestados desconfiaba de la base científica de las encuestas, y en particular pensaban que una muestra de sólo 1.000 ó 2.000 personas no podía representar correctamente a más de 200 millones de estadounidenses.¹

No tengo noticia de que nadie en España haya realizado un ejercicio parecido pero tengo la sospecha de que los resultados no serían muy distintos. ¿Y qué sucedería si hiciéramos esa encuesta sobre las encuestas a los periodistas? A juzgar por la frecuencia con la que se refieren a ellas, ha-

bría que suponer que sí creen en su validez. A modo de ilustración: una búsqueda rápida en Internet me dice que en el año 2005 se publicaron 1.097 noticias que contenían la palabra encuesta en *El País*, 1.442 en *El Mundo*, 1.055 en *Abc*, y 504 en *El Periódico*. Resulta una media de 2,8 noticias diarias por cada medio.

Sin embargo, es también absolutamente habitual encontrar en los medios de comunicación dudas, sospechas o descalificaciones sobre la validez de las encuestas. Y mi impresión anecdótica, cuando he charlado sobre este asunto con periodistas, es que muchos tienen una actitud que no difiere mucho de la de los encuestados de Gallup: se las creen (y las publican) pero sin estar muy seguros de por qué deben hacerlo, ni de cuándo una encuesta está bien hecha y es creíble y cuándo no.

Creo que esto es preocupante, pero tiene fácil remedio. En efecto, si

Josu Mezo es profesor de la Universidad de Castilla-La Mancha y editor de *Malaprensa.com* (editor@malaprensa.com).

los periodistas no conocen al menos unas nociones básicas de la ciencia y la técnica de las encuestas, no sabrán distinguir las buenas de las malas, o de las regulares, y cometerán errores graves en sus informaciones. Pero afortunadamente, la ciencia de las encuestas no es nada misteriosa. Sus principios básicos pueden ser entendidos por cualquiera, y no hace falta especial formación matemática o estadística. En estas páginas voy a intentar responder a algunas de las dudas más comunes al respecto, en particular respecto al tamaño de la muestra y su representatividad. Dejo para otra ocasión otros temas también importantes, como la redacción de las preguntas o la diferencia entre encuestas presenciales o telefónicas.

Para empezar, me gustaría responder a la pregunta más crucial de las que se pueden hacer al enfrentarse a las encuestas: ¿cómo es posible saber la opinión de millones de personas entrevistando sólo a unos pocos cientos o miles?

Mi impresión (de nuevo anecdótica) es que la mayor parte de la gente (y quizá de los periodistas) piensan que en realidad no es posible, pero no se atreven a decirlo alto y claro, debido al uso tan generalizado de las encuestas. Por lo tanto, siguen la corriente y aceptan que se debe de poder hacer. Pero como no entienden por qué, no saben cuáles son los requisitos que debe cumplir una encuesta para que esa proyección de los resultados sea

posible, y así uno acaba encontrando que se publican y difunden, casi con igual relieve noticias relacionadas con encuestas verdaderamente representativas y noticias relacionadas con encuestas que no tienen ningún valor.

En realidad, sí es cierto que se puede con una muestra muy pequeña conocer la opinión de una población muy grande, pero sólo si la muestra tiene una mezcla de componentes que se parece al total. A eso lo llamamos una muestra representativa, es decir, una muestra que tiene aproximadamente el mismo porcentaje que la población de cada grupo de personas que podamos imaginar: jóvenes y viejos, ricos y pobres, hombres y mujeres, religiosos o agnósticos, de izquierdas o derechas, licenciados o sin estudios, deportistas o sedentarios, casados o solteros... La manera ideal de obtener una muestra representativa es el puro azar, un sorteo, en el que todas las personas que forman parte de la población estudiada tengan la misma probabilidad de aparecer en la muestra. La ejecución práctica de este principio ideal es complicada, como veremos luego, pero de momento lo importante es quedarse con esta idea clave.

Esta noción esencial se puede demostrar con unos razonamientos sencillos. Supongamos que tenemos una población de un millón de personas, de las que el 60% están a favor y el 40% en contra de una determinada ley. Ponemos en una gran urna una

bola blanca por cada persona favorable y una bola negra por cada persona contraria a la ley. Revolvemos muy bien la urna, de forma que las bolas estén perfectamente mezcladas. Sabemos qué probabilidad hay de sacar una bola blanca (60%) y qué probabilidad hay de sacar una bola negra (40%). Por tanto, se pueden calcular también las probabilidades de diferentes combinaciones. Así, si sacamos dos bolas, la probabilidad de sacar dos bolas blancas es del 36%, una bola blanca y otra negra el 48%, y ninguna blanca y dos bolas negras el 16%. Si tomamos las dos bolas como una muestra del contenido de la urna, podríamos decir que la mitad de las veces (48%) la muestra sería aproximadamente correcta (una blanca y una negra), y la otra mitad de las veces (52%) sería totalmente incorrecta (dos blancas o dos negras).

¿Qué pasaría si extrajéramos una muestra de 4 bolas? Las probabilidades serían las siguientes: saldrían todas blancas un 13% de las veces; 3 blancas un 35%; 2 blancas también un 35%; sólo una blanca el 15% y ninguna blanca el 3%. Si agrupamos esos resultados entre los que dan un resultado aproximadamente correcto (2 ó 3 bolas blancas) y los que lo dan más bien incorrecto (0, 1 ó 4 bolas blancas), obtenemos que en el 69% de los casos nuestra muestra de cuatro bolas daría un resultado aceptable, y el 31% daría un resultado malo.

¿Y si sacamos muestras de 10 bo-

las? Para no marearles con los números, se los daré ya agrupados. Un 67% de las veces saldría un resultado que podríamos llamar bueno, que no se apartaría de la realidad de la urna más de un 10% (5, 6 ó 7 bolas blancas); un 23% de las veces saldría un resultado mediocre, con una desviación entre el 10 y el 20% (4 u 8 bolas blancas); sólo un 10% de las veces saldría un resultado malo, con un error de más del 20% (0, 1, 2, 3, 9 ó 10 bolas blancas).

Observemos la progresión tan rápida en precisión que hemos experimentado, al pasar de sacar dos bolas de la urna a sacar diez. Si consideramos resultados malos los que nos dan un error de más del 20%, tenemos que con dos bolas, serían malas el 52% de las muestras, con cuatro bolas el 31%, y con diez bolas, sólo un 10% de las muestras. Fíjense que lo que estaríamos haciendo, al extraer 10 bolas, sería el equivalente a entrevistar a diez personas al azar, de una población de un millón, y preguntarles si están a favor o en contra de una ley. El 67% de las veces que lo hiciéramos no nos equivocariáramos en más del 10%, el 90% de las veces obtendríamos un resultado que no se apartaría de la realidad más de un 20%, y sólo el 10% de las veces obtendríamos un resultado más alejado de la realidad. Naturalmente, acertar con un margen de error de $\pm 20\%$ cuánta gente está a favor o en contra de una ley, no parece gran cosa. ¡Pero es que

lo estaríamos haciendo con sólo diez entrevistados de una población de un millón!

Los mismos cálculos se pueden repetir para muestras de cualquier tamaño. Y así se puede calcular, usando las definiciones anteriores, que si sacáramos muestras de 50 bolas, el 89% serían buenas, el 11% mediocres y menos del 1% malas. Con 100 bolas, habría un 96% de las muestras buenas, un 4% mediocres y un 0,004% (4 de cada 100.000 muestras) malas.

Visto de otra forma, si el 96% de las muestras posibles tiene un error de hasta $\pm 10\%$, cuando saquemos una sola muestra sabemos que tenemos el 96% de probabilidades de que el error no sea mayor de $\pm 10\%$. Este es el famoso margen de error del que hablan las fichas técnicas de las encuestas. Lo que la ficha técnica nos dice es: si hiciéramos una muestra como esta muchas veces, el 95,5% de las veces (o el 95%, o el 99%) el error no sería mayor que $\pm X$ (la cifra que aparezca en la ficha). Por lo tanto, podemos de-

cir con un 95,5% de confianza que en esta muestra que hemos hecho el error no será mayor que X.

La tabla a pie de página presenta cuál es el margen de error máximo, en el 95,5% de las muestras, según el tamaño de la muestra y el porcentaje de personas (o de bolas) que tienen la característica que estamos estudiando.

Puede verse que para cada tamaño de la muestra (cada fila) el margen de error va creciendo a medida que crece la proporción en la población que tiene la característica medida, pero sólo hasta que esa proporción llega al 50%. A partir de ahí vuelve a disminuir, y de hecho el margen es el mismo para 60% que para 40%, y en general para cualquier par de valores que sumen 100. Como en las encuestas hay muchas preguntas, con diferentes proporciones en las diferentes respuestas, la ficha técnica suele dar el dato peor, el que corresponde al 50% (expresado en proporciones suele decir algo así como “cuando $p=q=0,5$ ”).

Margen de error, con un 95,5% de confianza

TAMAÑO DE LA MUESTRA	PORCENTAJE EN LA POBLACIÓN								
	10	20	30	40	50	60	70	80	90
100	6,0	8,0	9,2	9,8	10,0	9,8	9,2	8,0	6,0
300	3,5	4,6	5,3	5,7	5,8	5,7	5,3	4,6	3,5
500	2,7	3,6	4,1	4,4	4,5	4,4	4,1	3,6	2,7
800	2,1	2,8	3,2	3,5	3,5	3,5	3,2	2,8	2,1
1.000	1,9	2,5	2,9	3,1	3,2	3,1	2,9	2,5	1,9
1.500	1,5	2,1	2,4	2,5	2,6	2,5	2,4	2,1	1,5
2.000	1,3	1,8	2,0	2,2	2,2	2,2	2,0	1,8	1,3
2.500	1,2	1,6	1,8	2,0	2,0	2,0	1,8	1,6	1,2

Si me han seguido hasta aquí (espero que sí), ya saben todo lo esencial para comprender el milagro de que entrevistando a 1.000 personas podamos hablar de lo que piensan millones. Lo podemos hacer porque las matemáticas nos permiten calcular que si hiciéramos eso muchísimas veces, en el 95,5% de ellas obtendríamos en nuestra muestra un porcentaje de respuestas que no diferiría más del 3,2% del que nos daría toda la población, si pudiéramos entrevistarla.

Pero atención, el milagro de que entrevistando a unos pocos cientos o miles podamos estimar la opinión de millones depende esencialmente de que la elección de los miembros de la muestra haya sido aleatoria. Eso es lo que nos permite calcular el margen de error. Si ese criterio esencial no se cumple, la encuesta no sirve para prácticamente nada, por muy grande que sea su muestra, porque si no hay aleatoriedad, los encuestados pueden provenir desproporcionadamente de grupos sociales peculiares, por el lugar de residencia, por la ideología, por sus intereses... y producir unos resultados totalmente distintos a los que se obtendrían estudiando a toda la población.

El ejemplo real más célebre que ilustra esta verdad se produjo en 1936 cuando la revista *Literary Digest* envió una encuesta electoral a 10 millones de personas, basándose en su propia lista de suscriptores, y los registros telefónicos y de coches matriculados.


Los 2,4 millones de respuestas recibidas le llevaron a anunciar la victoria de Landon sobre Roosevelt, por 57% contra 43%. George Gallup criticaba esa forma de encuestar, pues tenía sesgos importantes favorables a los más ricos. En su lugar, él hizo una encuesta a una muestra representativa de 5.000 personas, y predijo una victoria de Roosevelt por 56% contra 44%. El resultado real fue una victoria de Roosevelt, aún mayor (61% contra 37%), que lanzó a la fama a Gallup y dejó claramente establecida, no sólo en la teoría, sino en la práctica, la superioridad de las buenas muestras (aunque pequeñas) sobre las malas muestras (por grandes que sean).²

La pregunta más importante, por tanto, que hay que tener siempre presente antes de leer (y de contar a otros) los resultados de una encuesta es: ¿cómo han sido seleccionados los entrevistados? ¿Por un procedimiento aleatorio o de otra forma? Si la encuesta se ha realizado con una muestra aleatoria, podemos seguir adelante. Si no se sabe cómo se ha seleccionado la muestra, o se sabe con seguridad que la selección no ha sido aleatoria, sus datos no pueden ser el fundamento de ninguna noticia con un mínimo de rigor, porque no es posible saber si lo que opinan los encuestados se parece mucho, poco, o nada a lo que opinaría el conjunto de la población.

Ahora bien, como antes he apuntado, el principio ideal de la muestra puramente aleatoria raramente se

puede llevar a cabo. Hay dos razones por las que las muestras que realmente se hacen en investigación social no corresponden exactamente a las muestras aleatorias puras de las que he venido hablando hasta ahora. La primera razón es que no tenemos un registro con los nombres y los datos de contacto de todos los españoles, o mejor dicho, los registros que existen no se pueden usar, por razones legales, para hacer encuestas. Por esa razón, aunque se quiera hacer una muestra aleatoria, no se puede hacer de manera directa, sino indirecta, combinando el sorteo con las cuotas. Así, por ejemplo, se divide la muestra proporcionalmente entre las provincias, y luego entre los municipios. Los más grandes reciben una cuota directa, y los demás se agrupan por tamaños (hasta 1.000 habitantes, de 1.000 a 5.000, de 5.000 a 20.000... o similar) y se adjudica a cada grupo de municipios las encuestas que les correspondan, según su población. Todo esto garantiza que la muestra mantiene correctamente las cuotas territoriales y por tipo de población.

A partir de ahí, se realizan una serie de sorteos. Se decide por sorteo, entre los municipios pequeños, en cuál o cuáles se va a hacer la encuesta. Lue-



El rigor de una encuesta depende de cómo se hayan seleccionado los entrevistados.

go se sortea, en todos los municipios incluidos en la muestra, los domicilios a los que contactar. Si la encuesta es telefónica, se introducen en el ordenador todos los números del municipio y se llama aleatoriamente. Si la encuesta es presencial, se sigue un sistema que va enviando a los encuestadores por las calles y luego por las casas, siguiendo rutas aleatorias. Esto se hace a diferentes horas del día, para evitar sesgos según

quien esté o no en casa. Aun así, hay grupos que son más difíciles de encontrar (típicamente los varones ocupados), pero esto tiene una solución de la que hablaré enseguida.

Llegados a este punto, aparece un segundo problema: a diferencia de las bolas de la urna, que no pueden negarse a enseñarnos su color, las personas elegidas por un sorteo pueden negarse a responder a una encuesta. De hecho, son muchas las personas que se niegan a participar en las encuestas. En encuestas telefónicas en España, aproximadamente, dos de cada tres personas contactadas se niegan a participar o interrumpen la encuesta una vez iniciada. En ese caso nos vemos obligados a sustituir a la persona que se niega por otra, pero hay razones para sospechar que el perfil demográfico de las personas

que no responden es distinto (por edad, por estudios, por ocupación, por género) del de las que aceptan responder. Para resolver este problema, y el mencionado antes de los grupos sociales difíciles de encontrar en casa, se vuelve a introducir un sistema de cuotas: puesto que sabemos aproximadamente la composición demográfica del conjunto de la población, por edad, sexo, ocupación y estudios, al realizar la encuesta se controla que la composición de la muestra por esas variables no sea muy diferente de la de la población total.

Prácticamente todas las empresas de investigación de opinión pública serias, en España y en otros países, realizan variantes de este procedimiento, combinando aleatoriedad (aplicada a los domicilios, no a las personas) y cuotas (primero territoriales, y luego por grupos demográficos). Si se hace con cuidado hay pruebas suficientes de que se pueden obtener resultados que no se alejan apenas de los que se obtendrían con muestras aleatorias puras, en las que todo el mundo estuviera obligado a contestar, como las bolas de una urna. Por eso aceptamos estas muestras como si fueran aleatorias, y que en las fichas técnicas que las acompañan se calculan los márgenes de error como los de la tabla de arriba.

¿Y cuáles son las encuestas con muestras no aleatorias? Pues, naturalmente, no lo suelen proclamar. Pero se pueden detectar casi siempre

con poco esfuerzo. Un primer tipo muy común son las muestras que llamamos de conveniencia. Esto es, muestras hechas con quienes se tiene a mano. Por ejemplo, una asociación encuesta a sus miembros (y seguramente no a una muestra aleatoria de los mismos), y luego difunde los resultados como si sus opiniones fueran representativas de todo el sector profesional o social al que la asociación pretende representar. O una empresa hace una encuesta entre sus consumidores, y obtiene conclusiones que generaliza a todos los consumidores de ese tipo de productos, o a toda la población. Aunque no lo anuncien en titulares, no será seguramente difícil encontrar en la nota de prensa correspondiente la información que indique que estamos ante una muestra de este tipo. Estas encuestas pueden servir, como mucho, y según los casos, como pistas, indicios o apuntes muy borrosos del estado de la opinión pública, teniendo en cuenta siempre, según quién y cómo haya hecho la encuesta, los sesgos de la muestra. Son tantas las cautelas que hay que tomar, de todos modos, que parece preferible no construir nunca una noticia centrada en una encuesta de este tipo.

Por otro lado, la muestra no aleatoria más común actualmente es la muestra en la que los participantes se seleccionan a sí mismos, porque se encuentran con la encuesta de una forma no aleatoria y deciden contes-

tar en función de su interés por el tema. Este el caso de (prácticamente) todas las encuestas *online*, colocadas en páginas web, para que los visitantes expresen su opinión. Dejando a un lado el problema, ya muy grave, de que no es posible evitar que una misma persona responda varias veces (aunque sí dificultarlo), la razón esencial por la que esas encuestas no sirven para nada fiable es que la selección de la muestra no es aleatoria. Las personas que navegan por Internet no son una muestra aleatoria de la población: son desproporcionadamente más jóvenes, más ricos, con más estudios, y varones. Los que entran en tal o cual página, a su vez, no son una muestra aleatoria: según la página de que se trate, su ideología, intereses, edad, estudios... serán distintos. Finalmente, los que ven la pregunta colocada en la página y se sienten motivados para participar no son tampoco una muestra aleatoria, ya que serán personas con más interés por la cuestión en particular que la media de la población.

Todo esto debería ser bien conocido. Y sin embargo una vez tras otra, más de un medio y más de dos dan aire a encuestas hechas así. En el último año he podido leer, por ejem-

Las encuestas *online* nunca pueden ser fiables.

plo, en los principales periódicos españoles, y en titulares, que el 65% de los españoles practican sexo en el coche, que los trabajadores norteamericanos pierden una media de dos horas al día en Internet, que tres de cada cinco farmacéuticos están a favor del cannabis terapéutico, y que el 20% de los europeos se queda en la oficina para usar Internet. Todos estos datos provenían de encuestas *online* con muestra autoseleccionada, y por tanto, básicamente inservible.

Esto nos lleva a un asunto delicado: las encuestas en las páginas web de los periódicos, radios, etc... He hecho un pequeño estudio (rápido y seguramente impreciso) en el que he llegado a las siguientes conclusiones: de los 10 periódicos españoles de pago de información general con más tirada, 8 incluyen en su página web (6 en la portada) algún tipo de encuesta (las excepciones son *La Razón* y *El Correo*). Por alguna razón, este entusiasmo por la encuesta *online* es mayor en España que en otros países. En una cata rápida de los cuatro periódicos de calidad más vendidos en EEUU, Gran Bretaña, Francia e Italia (posiblemente con errores), he encontrado que sólo 5 de los 16 publican encuestas en su página (uno de ocho en los medios

en inglés, y cuatro de ocho en Francia e Italia).

Algunas de estas encuestas tienen un éxito notable, con miles y a veces decenas de miles de respuestas (he llegado a ver en periódicos españoles más de 100.000). Y sin embargo, por las razones que he explicado más arriba, no tienen valor alguno como medidas ni de la opinión pública general, ni de los lectores de ese periódico, ni siquiera, realmente, de los lectores *online* de ese periódico. Quiero subrayar, una vez más, que el número de encuestados, por muy alto que sea, no puede compensar por una muestra no aleatoria. Siempre es mejor una muestra pequeña, y aleatoria, que una muestra muy grande, pero no aleatoria.

El valor de estas encuestas *online*, por tanto, va poco más allá del puro entretenimiento. Como tal, quizá, pueden ser aceptables, si van acompañadas de una nota explicativa que indique a los lectores que los resultados no tienen validez científica alguna. Sin embargo, sólo dos de los ocho periódicos más vendidos que incluyen encuestas en sus páginas hacen este tipo de advertencias (son *El País* y *El Diario Vasco*). Los otros seis permiten que sus lectores vean los resultados sin ninguna advertencia de su total falta de representatividad. Algún periódico va incluso más lejos y tras pasa a su edición en papel, de manera regular, resultados de sus encuestas *online*, de nuevo sin advertencia al-

guna sobre la total falta de significación de esos resultados.

En conclusión, la regla número 1 sobre las encuestas es la referida al modo de selección de la muestra. En su versión más simple: muestra aleatoria (o con combinación de aleatoriedad y cuotas), vale; muestra no aleatoria, no vale.

A partir de ahí, podemos hacernos algunas preguntas adicionales. En primer lugar, ¿es mejor la encuesta cuanto mayor sea el número de encuestados? Hasta cierto punto esto es verdad, pero llega un momento en el que añadir más gente a la muestra no aporta gran cosa. En la tabla de la página 177 puede verse que, para cualquier proporción en la población (cada columna), el margen de error va siendo menor cuanto mayor es la muestra. Pero la disminución no es uniforme, ni el error proporcional al número de entrevistados. Los primeros cientos de entrevistados disminuyen muchísimo el margen de error, y luego cada vez cuesta más ir mejorando la precisión. Así, por ejemplo, cuando la proporción es del 50%, pasar de 100 a 500 entrevistados reduce el error un 5,5%; de 500 a 1.000 un 1,3%; de 1.000 a 1.500 un 0,6%; de 1.500 a 2.000 un 0,4%, y de 2.000 a 2.500 un 0,2%. Podemos decir, tomando un término de la economía, que hay un rendimiento marginal decreciente, por el que las personas que vamos añadiendo a la muestra (con notable gasto) nos aportan cada vez me-

nor mejoría de la precisión.

¿Cuál es el tamaño ideal, entonces, el equilibrio correcto entre precisión y esfuerzo económico y de otros recursos para conseguirla? Convencionalmente suelen usarse en estudios sociales y políticos muestras de unas 1.000 personas, que tienen un margen de error de $\pm 3,2\%$, para un nivel de confianza del 95,5%, en el caso más desfavorable de que la variable medida la tenga un 50% de la población. Esto significa en la práctica que si un 50% de las personas entrevistadas dicen que tienen una determinada opinión, podemos estar al 95,5% seguros de que entre la población los que tienen esa opinión estarán en una horquilla entre 46,8% y 53,2%. Como hemos visto, el margen de error es algo menor cuando el porcentaje es distinto al 50%, pero a la hora de diseñar la encuesta, como los porcentajes de las respuestas son desconocidos, los investigadores se ponen en lo peor y calculan el número de entrevistados para el caso más desfavorable.

En general se acepta, como he dicho, que un margen de $\pm 3,2\%$ (una horquilla de 6,4%, con un 95,5% de confianza), es suficiente en la mayor parte de los casos. Si queremos reducir el margen de error a $\pm 2\%$ (una hor-

Una muestra pequeña, y aleatoria, siempre es mejor que una muestra muy grande, pero no aleatoria.

quilla de cuatro puntos), tendríamos que entrevistar a 2.500 personas, más del doble. Y bajar el margen a $\pm 1\%$ es ya prohibitivo: serían necesarias 10.000 entrevistas. En cambio, si nos podemos conformar con un margen de error de $\pm 4\%$ (una horquilla de ocho puntos), nos bastaría con entrevistar a 625 personas. Y bastarían 400 si estamos dispuestos a aceptar un margen de $\pm 5\%$ (horquilla de diez puntos).

Esta sería la regla número 2: no hay un número perfecto de entrevistados para una encuesta. En líneas generales se acepta como suficiente una muestra de 1.000 personas, pero según los propósitos de la encuesta, pueden ser suficientes 400 personas, o ser necesarias 2.500.

Una tercera duda, menor, que conviene despejar es la siguiente: ¿el número de personas entrevistadas para que la encuesta sea fiable es de alguna manera proporcional a la población a que se refiere? Esto es, para conseguir una encuesta igualmente fiable, ¿es necesario entrevistar a más personas en Madrid (6 millones de habitantes) que en Navarra (0,5 millones) y más en Europa (350 millones) que en España (44 millones)?

La respuesta está implícita en la tabla anterior, en la que no se hacía men-

ción al tamaño de la población. Pien- sen en el ejemplo de las bolas blancas y rojas. Decíamos que podemos calcu- lar la probabilidad de sacar una bola blanca (60%) y negra (40%). Y a partir de ahí calculamos la probabilidad de las combinaciones (muestras) de 2, 4, 10, 100 ó 500 bolas. Estrictamente ha- blando, la probabilidad de sacar una bola blanca o negra sólo es del 60% y el 40% antes de sacar la primera. Lue- go, va cambiando con cada extracción. Si la primera bola es blanca y la urna tiene un millón de bolas, en la siguien- te extracción las probabilidades son de 599.999/999.999 para las bolas blan- cas y 400.000/999.999 para las negras. Pero si en la urna hubiera 100.000 bo- las, las probabilidades serían respec- tivamente de 59.999/99.999 y 40.000/ 99.999. En cada nueva extracción ha- bría que recalcular las probabilidades, y esas probabilidades serían ligerísimamente diferentes con una urna de 100.000 o de un millón o de 40 millo- nes de bolas. Sin embargo, si la urna es suficientemente grande (mayor de unas 30.000 personas), esta diferencia es tan nimia que a efectos prácticos se puede ignorar, de forma que, como en la tabla vista más arriba, se calcu- la el margen de error de las encues- tas sin tener en cuenta el tamaño de la población de la que se obtiene la muestra.

La única excepción se produce cuando la proporción entre la mues- tra y la población es muy alta. Si en- trevistamos a 1.000 personas en un

pueblo de 5.000 personas, o en una empresa con 3.000 empleados, los márgenes de error son menores que si lo hacemos en una ciudad de 100.000 habitantes. En todo caso, la ficha técnica del estudio habrá teni- do esta circunstancia en cuenta, y nos informará debidamente del margen de error correcto.

Esta sería la regla número 3: el mar- gen de error, o la precisión, de una encuesta de 1.000 personas es la misma en toda población grande (mayor de 30.000 personas). Es por tanto perfec- tamente legítimo hacer estudios com- parativos entre territorios con pobla- ciones muy distintas y usar muestras de igual tamaño. En realidad es justa- mente lo correcto, para que el margen de error sea en todos los casos similar.

Tenemos por tanto, en relación con la representatividad de las mues- tras, tres reglas de oro para interpre- tar (y contar a otros) encuestas. La pri- mera y más importante, con mucho, es que sólo las encuestas con mues- tra aleatoria (con los matices antes vistos) pueden servirnos para cono- cer al conjunto de la sociedad. Sólo encuestas de este tipo deberían cons- tituir la materia prima central para las noticias. Las encuestas que no cumplen este requisito esencial pue- den servir como material secundario, siempre con cautelas, o como entre- tenimiento (en particular las basadas en encuestas *online*). Superado este fil- tro esencial, todo lo demás es mucho menos importante. En particular (se-

gunda regla), el tamaño de la muestra es mucho menos relevante de lo que se cree. Una muestra no es proporcionalmente mejor cuanto más grande sea. Si la muestra es representativa, una muestra de 1.000 personas es suficiente para la mayoría de los propósitos, y esto es independiente (tercera regla) del tamaño de la población. Como usuarios de encuestas, además, las dos últimas reglas no nos deben preocupar mucho, ya que si la muestra es aleatoria, la ficha técnica del estudio nos informará del margen de error (el porcentaje que hay que sumar y restar a cada resultado para calcular la horquilla de datos probable en la población).

Como les prometí, no hay nada misterioso en la ciencia básica que está detrás de las encuestas. De ella se derivan reglas muy sencillas, que nos permiten identificar qué encuestas tienen validez, y nos permiten decir algo con rigor, y cuáles no. Así que ya saben, la próxima vez que se encuentren con una encuesta mala, pueden responder con confianza y aplomo a su jefe, o al gabinete de prensa de turno: “Esa encuesta no vale para nada y no pienso publicar una noticia basándome en ella”. ¿A que les apetece? Todos se lo agradeceremos. ➔

1 Newport, Frank et al. 1997. ‘How are polls conducted?’, en Where America Stands. John Wiley & Sons, Inc.: <http://media.gallup.com/PDF/FAQ/HowArePolls.pdf>

2 ‘Introduction to Polling’ en http://www.pages.drexel.edu/~pa34/ptn1.htm#_ftn1

Informe anual de la profesión periodística 2005

Pedro Farias (director) , 376 págs., 24 euros

El primer estudio con referencias estadísticas desde las perspectivas profesional, laboral, asociativa y formativa.

DE VENTA EN LA A.P.M.

Asociación de la Prensa de Madrid

A05

INFORME ANUAL DE LA
PROFESIÓN PERIODÍSTICA
2005