

¿Son verdaderas las encuestas? Derribando algunos mitos populares

Luis Sánchez

Director de estadística y muestreo
Mail: Luis.Sánchez@ipsos.com

Existen dos maneras de obtener información sobre una población: censando a todas las personas o seleccionando una muestra para luego inferir sobre toda la población. Las encuestas están relacionadas a esta segunda opción y en torno a ellas usualmente se tejen una serie de mitos que ponen en tela de juicio su credibilidad. Veamos qué tan ciertos o falsos son algunos de los más populares mitos asociados con ellas.

Un primer mito es “pocos casos no permiten tener buenos resultados”, lo cual es falso. Basta una gota de nuestra sangre para saber de nuestra salud, esta magia se llama “inferencia”, la cual consiste en sacar conclusiones sobre toda una población a partir de una muestra representativa. ¿Cómo funciona esto? Pues, los elementos son seleccionados aleatoriamente y las características que se observan en ella se pueden atribuir o inferir a la población objetivo (Scheaffer, Mendehall, Ott, 2006). Por ejemplo, para que con 1200 casos podamos tener resultados de Perú, nos aseguramos de que la muestra tenga casos en el ámbito urbano rural y las macro regiones (Lima, norte, centro, sur, oriente) y así sea representativa de todo el país.

Conceptos claves de muestreo



Población objetivo

Es la colección completa de observaciones que se desea estudiar. Su definición es clave y debe detallarse lo más posible. Por ejemplo hombres y mujeres de 15 a 50 años, de los niveles socioeconómicos de Lima Metropolitana.



Marco muestral

Es la lista de las unidades muestreado. Por ejemplo para las encuestas telefónicas es la relación de todos los números celulares, y para las encuestas en hogares el marco muestral es la cartografía del INEI.



Margen de error

Es el error de muestreo, el cual resulta al considerar una muestra y no al examinar a toda la población. Por ejemplo si en una encuesta se obtiene que 30% de limeños consume la gaseosa Pepito, y el margen de error es de 5%, significará que el verdadero valor de la población está entre 25% y 35%.

Un segundo mito popular es aquel que refiere que “las encuestas son falsas, a mí nunca me han entrevistado”. Ante ello deberíamos de preguntarnos, “la Tinka existe, pero ¿conoces algún ganador?”. Detrás de ello está el concepto de **“probabilidad”**, hay técnicas de muestreo no probabilísticas y probabilísticas, estas últimas hacen que la probabilidad de selección de las personas sea igual para todas (INEI, 2011). Por lo tanto, ni la Tinka ni las encuestas son falsas, sino que la probabilidad de selección en ambas es baja. Según el INEI somos 33 millones de habitantes en Perú, por lo que la probabilidad que uno sea seleccionado es de 0.00000003. Ipsos, por ejemplo, utiliza técnicas probabilísticas en encuestas nacionales, asegurando así que toda persona tenga chance de ser seleccionada.



Fuente: Elaboración propia

Finalmente, uno de los mitos que circulaba en la coyuntura de las elecciones fue que “desproporcionar la muestra da malos resultados”. Esto nuevamente es falso. La desproporción es usual y se hace con el fin de tener más precisión en algún segmento, reducir costos, etc., pero en estos casos la muestra debe pasar por un proceso de “ponderación”, que consiste en asignar pesos a los encuestados de forma que al hacer estimaciones se conserve la distribución poblacional (Kish, 1995). Veamos un ejemplo:

Ejemplo de ponderación

1	<p>De fuentes oficiales se tiene la distribución de personas por Lima y el Interior del país</p> <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 15%;"></td><td style="width: 15%; text-align: center;">Población</td><td style="width: 70%;"></td></tr> <tr> <td style="text-align: center;">Lima :</td><td style="text-align: center;">34.5%</td><td></td></tr> <tr> <td style="text-align: center;">Interior:</td><td style="text-align: center;">65.5%</td><td></td></tr> </table>		Población		Lima :	34.5%		Interior:	65.5%				
	Población												
Lima :	34.5%												
Interior:	65.5%												
2	<p>Suponiendo que la muestra es de 1200 casos, la muestra proporcional será como sigue:</p> <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 15%;"></td><td style="width: 15%; text-align: center;">Población</td><td style="width: 70%; text-align: center;">Muestra proporcional</td></tr> <tr> <td style="text-align: center;">Lima :</td><td style="text-align: center;">34.5%</td><td style="text-align: center;">414</td></tr> <tr> <td style="text-align: center;">Interior:</td><td style="text-align: center;">65.5%</td><td style="text-align: center;">786</td></tr> </table>		Población	Muestra proporcional	Lima :	34.5%	414	Interior:	65.5%	786			
	Población	Muestra proporcional											
Lima :	34.5%	414											
Interior:	65.5%	786											
3	<p>Pero si deseamos tener una mejor precisión para los resultados de Lima se eleva su muestra a 500 casos, y haciendo que ahora represente un 41.7%</p> <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 15%;"></td><td style="width: 15%; text-align: center;">Población</td><td style="width: 70%; text-align: center;">Muestra desproporcionada</td></tr> <tr> <td style="text-align: center;">Lima :</td><td style="text-align: center;">41.7%</td><td style="text-align: center;">500</td></tr> <tr> <td style="text-align: center;">Interior:</td><td style="text-align: center;">58.3%</td><td style="text-align: center;">700</td></tr> </table>		Población	Muestra desproporcionada	Lima :	41.7%	500	Interior:	58.3%	700			
	Población	Muestra desproporcionada											
Lima :	41.7%	500											
Interior:	58.3%	700											
4	<p>Para tener los resultados totales y corregir la desproporción de la muestra, se calcula los factores de ponderación, que se origina al dividir los %ideales (población) entre los %reales (muestra)</p> <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 15%;"></td><td style="width: 15%; text-align: center;">%Ideal</td><td style="width: 15%; text-align: center;">%Real</td><td style="width: 60%; text-align: center;">Pesos</td></tr> <tr> <td style="text-align: center;">Lima :</td><td style="text-align: center;">34.5%</td><td style="text-align: center;">41.7%</td><td style="text-align: center;">$34.5/41.7=0.83$</td></tr> <tr> <td style="text-align: center;">Interior:</td><td style="text-align: center;">65.5%</td><td style="text-align: center;">58.3%</td><td style="text-align: center;">$65.5/58.3=1.12$</td></tr> </table>		%Ideal	%Real	Pesos	Lima :	34.5%	41.7%	$34.5/41.7=0.83$	Interior:	65.5%	58.3%	$65.5/58.3=1.12$
	%Ideal	%Real	Pesos										
Lima :	34.5%	41.7%	$34.5/41.7=0.83$										
Interior:	65.5%	58.3%	$65.5/58.3=1.12$										
5	<p>Estos pesos son asignados a cada encuestado, de tal forma que al tener el resultado total la muestra tenga la misma distribución de la población, es decir:</p> <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 15%;"></td><td style="width: 15%; text-align: center;">Muestra Ponderada</td><td style="width: 70%;"></td></tr> <tr> <td style="text-align: center;">Lima :</td><td style="text-align: center;">34.5%</td><td></td></tr> <tr> <td style="text-align: center;">Interior:</td><td style="text-align: center;">65.5%</td><td></td></tr> </table>		Muestra Ponderada		Lima :	34.5%		Interior:	65.5%				
	Muestra Ponderada												
Lima :	34.5%												
Interior:	65.5%												

Fuente: Elaboración propia

En conclusión, las encuestas son verdaderas y funcionan bien gracias a la estadística inferencial, una ciencia que es parte fundamental de la vida de las personas; por ello conocer sus principios es relevante para saber qué resultados de encuestas son confiables y cuáles no.

Siempre que tenga el resultado de una encuesta entre manos y quiera validar que tan confiable es o no, revise la ficha técnica, valide que tenga un buen diseño y confirme que haya sido trabajada por expertos en el tema.

Fuentes:

- Lohr, S. (2019). Sampling: Design and analysis. Arizona State University: Chapman and Hall/CRC.
- Instituto Nacional de Estadística e Informática – INEI (2011). Buenas prácticas de una encuesta por muestreo. Perú.
- Scheaffer, R., Mendehall, W. y Ott,R. (2006). Elementos de muestreo. Madrid, España: Paraninfo Cengage Learning.
- Kish, L. (1995). Survey sampling. New York, Estados Unidos: John Wiley & Sons.