



Para leer encuestas (electorales)

Carlos Ponce

Director de proyectos – Ipsos Public Affairs Perú

E-mail: Carlos.Ponce@Ipsos.com

En época electoral, la necesidad por entender la estadística, las probabilidades y los números en general es apremiante, ya que vivimos en medio de un bombardeo de información que se vuelve muy difícil de digerir y que, antes de ayudarnos a formar una opinión, suele confundir a la mayoría.

En este sentido, es necesario revisar algunas cosas que debemos tomar en cuenta cuando queremos entender las encuestas que vemos, oímos o leemos en los medios de comunicación.

Muestra y universo: el mito del “a mí nunca me han encuestado”

Muchos detractores de la investigación cuantitativa suelen usar este argumento para desacreditar la metodología iniciada y probada, hace más de 70 años, por el periodista e investigador George Gallup, quien en 1936 logró predecir los resultados de las elecciones presidenciales en Estados Unidos usando una muestra de 5 mil casos entre la población electoral estadounidense, que en ese entonces superaba los 45 millones. La revista *Digest* había recogido información de más de dos millones de personas, sin embargo, no le atinó al ganador.

La idea planteada por Gallup es que no tienes que entrevistar a todo el conjunto de personas que quieres investigar (universo) para obtener una estimación confiable; basta con elegir a un subgrupo representativo de individuos (muestra), siguiendo técnicas estadísticas. Lo que garantiza la validez de este método es **la aleatoriedad** en las etapas del muestreo y que los individuos del universo tengan **alguna probabilidad conocida de ser elegidos**.

Es así que, con una muestra, podemos estimar lo que piensan los cerca de seis millones de electores limeños con un margen de error conocido. Aumentando la muestra reduces el margen de error, sin embargo, llega un punto en el que el aumento de la muestra no reduce considerablemente este margen. **Una muestra más grande no garantiza un mejor resultado**, solo un costo mayor y tiempos poco manejables para cualquier investigación. Si ves una encuesta de 10 mil casos en Lima, no es mejor que una de 500, lo más probable es que sea falsa o no haya seguido métodos científicos en el muestreo.



Para estimaciones en Lima se trabajan muestras de entre 500 y 1000 casos, para nivel nacional muestras de 1000 a 2000 casos, ¿se entiende entonces por qué es sumamente probable que millones de personas no hayan sido encuestadas?

Diferencia entre sondeo y encuesta realizada con una muestra

Hay muchos métodos para recoger información sobre una población, pero no todos son científicos. Los métodos no científicos son conocidos en el Perú como sondeos. ¿Por qué decimos que no siguen un método científico? **Porque que no tienen aleatoriedad** y porque **no todos los individuos del universo que pretenden representar tienen alguna probabilidad de ser elegidos**.

Ejemplos claros de sondeos son los que se hacen *online* en webs de diarios, o vía telefónica en programas televisivos. La limitación de estos sondeos es que no pueden extrapolarse al total de la población, solo representan a los lectores del diario que decidieron participar en la encuesta *online*, o a los televidentes que decidieron llamar al programa y entró su llamada.

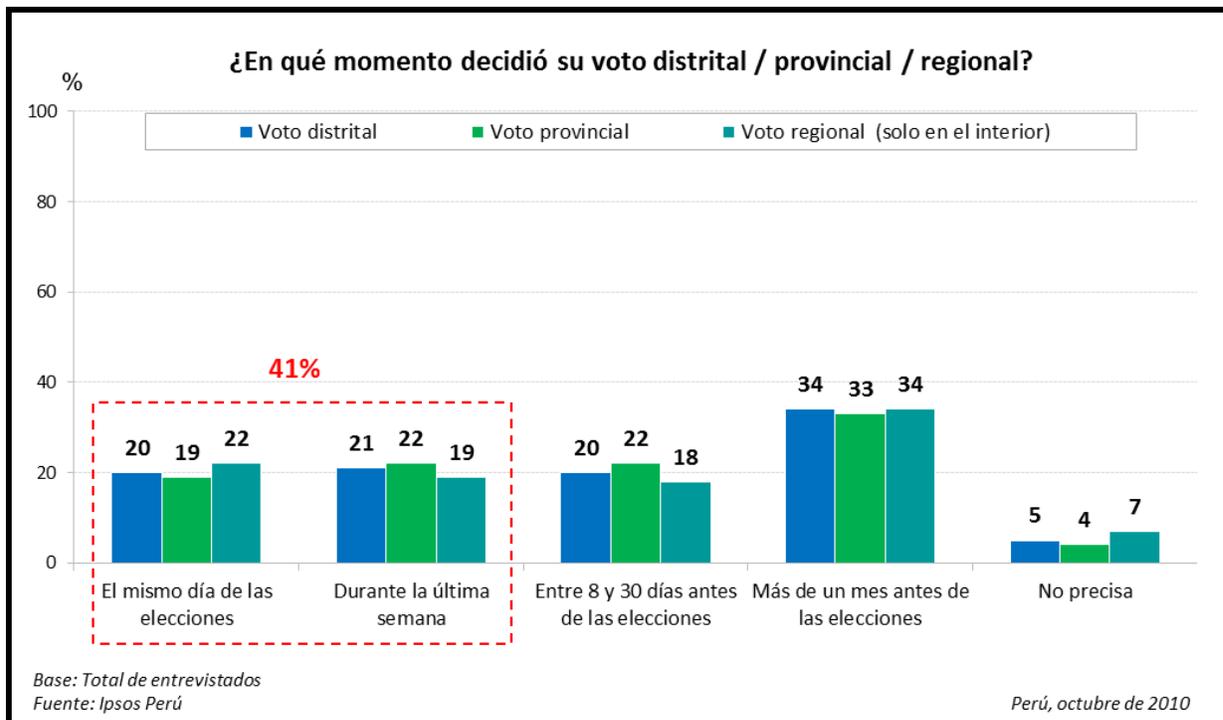
Es claro que en estos casos no puedes garantizar que te llamen personas de todas las posturas o clases sociales; es probable que te llamen solo los que votan por un candidato, como es probable que una encuesta *online* sea mayormente contestada por jóvenes con acceso a internet.

Comparación entre encuestas: distintas empresas, distintas fechas, distintas metodologías

Es muy común que los medios de comunicación hablen de subidas y bajadas de puntos en la aprobación presidencial o en la intención de voto de un candidato. En este caso la clave es ver a qué números están haciendo referencia para decir si algún porcentaje bajó o subió.

Estrictamente se pueden comparar encuestas que hayan tenido el mismo método de muestreo, la misma muestra, preguntas redactadas del mismo modo y con las mismas opciones de respuesta. Por eso lo ideal sería no mezclar datos de encuestadoras distintas al hacer la comparación, ya que pueden tener métodos de muestreo diferentes, métodos de selección de los entrevistados distintos o fraseos disímiles en las preguntas.

No obstante, cuando se asume la científicidad y seriedad en dos investigaciones distintas, se puede tomar ciertas licencias para hacer comparaciones. En estos casos lo que hay que evaluar son las fechas de aplicación, ya que la opinión pública puede ser muy variable; prueba de esto son las últimas elecciones presidenciales, las municipales en Lima o la revocatoria. En esta última, 25% decidió su voto en la última semana (semana en que publicar encuestas está prohibido), y en las regionales, provinciales y distritales de 2010 fue un 40% el que decidió su voto la última semana.



Ver más tendencias que variaciones de puntos porcentuales: el margen de error

Al leer una variable que ha sido *trackeada* en varias mediciones, debemos ver si la variación con respecto a la medición anterior es mayor al margen de error.

Por ejemplo, si tienes un margen de error de 3% en la encuesta nacional urbana que evalúa la gestión del presidente Humala, y la aprobación pasa de 27% a 24%, no podemos hablar de una caída en la aprobación. El primer dato puede ser 30%, 27% o 24%, y el segundo podría ser 27%, 24% o 21%.

Lo mismo ocurre en las encuestas de intención de voto, cuando la diferencia entre dos candidatos no supera el margen de error se habla de un empate técnico.

Tipo de recojo de información: encuesta de intención de voto vs. cédula

Hay dos datos más que suman a la confusión colectiva: la intención de voto sugerida y la simulación del voto con cédula.



En la primera, se le pregunta a la persona por quién votaría si las elecciones fueran mañana, y se le ofrece una tarjeta con los candidatos a medir. En el caso de la simulación con cédula, se entrega al encuestado una cédula como la que tendrá en frente el elector el día de las elecciones, la persona marca de modo secreto y deposita su voto en un ánfora.

Al analizar estas dos técnicas, hay que tomar en cuenta los factores que pueden hacer diferir el dato de intención de voto del dato de la cédula. En la pregunta sugerida se muestra el nombre del candidato y del partido, en la cédula provincial y distrital solo figura el nombre del partido y el símbolo. Es decir, una persona puede declarar que va a votar por un candidato, pero si no sabe el símbolo, eso se verá reflejado en la cédula, donde no sabrá qué marcar.

En este mismo sentido, la gente puede no conocer a un candidato distrital, pero si marca un símbolo conocido en la provincial, hay una probabilidad de que marque el mismo en la distrital, es lo que se suele llamar "arrastré".

Es por esto que la medición más efectiva es la cédula, ya que simula la situación real a la que se enfrentarán los electores el día mismo de la votación.

Votos emitidos vs. votos válidos

De la simulación con cédula se presentan dos datos: votos emitidos y votos válidos, tal como se presentan los datos el día de las elecciones.

Los votos emitidos salen del conteo del total de las cédulas, y en sus resultados se incluye los que marcaron blanco o viciado. Los votos válidos se obtienen generando un nuevo 100% solo de votos válidos, y excluyendo las cédulas en blanco y las viciadas.

Boca de urna y conteo rápido

Finalmente, el día de las elecciones suelen publicarse dos datos: el flash con los datos de la boca de urna, hacia la mitad de la tarde, cuando cierran las votaciones, y el conteo rápido, hacia el final del día con el conteo que van haciendo las mesas de sufragio.

La boca de urna es un método no probabilístico, y su científicidad es cuestionada por muchos, sin embargo, ha demostrado ser útil en ciertas coyunturas. La metodología consiste en muestrear los colegios electorales según la cantidad de población que alberguen, y se escogen los lugares de votación que representen mejor a la población en estudio. Luego, los encuestadores se ubicarán afuera de los lugares de votación, y aplicando saltos sistemáticos para elegir a los entrevistados (intercalando encuestados según género y edad, y dejando un tiempo determinado entre encuestado y encuestado) se acercarán a las personas que ya hayan votado para que repliquen su voto.



Estos datos eran antes transmitidos vía telefónica por nuestros encuestadores hacia una central donde se digitaban los datos y se ingresaban al sistema que nos permitía procesarlos; actualmente, la tecnología nos permite transmitirlos en tiempo real con una *Tablet* hacia un servidor, que los almacena para luego ser procesados.

El conteo rápido es distinto, no hay encuestados. Los encuestadores se quedan en los lugares de votación a esperar el conteo de los miembros de mesa. Como ya lo mencionamos, estos lugares de votación han sido previamente seleccionados como representativos por la cantidad de población que vota en ellos. Se espera que el conteo se publique, y se transmite a nuestra central. En este sentido, este es el dato más cercano al que otorgará la ONPE luego de escutar todas las actas.

Como vemos, es fácil confundirse en esta maraña de datos. Por eso se hace necesario entender los números que los medios nos presentan, cuestionarlos y analizarlos a fondo, si es que queremos tener una visión acertada de lo que las técnicas cuantitativas nos ofrecen, y no caer en el uso truculento que algunos hacen con fines de propaganda política.