

Las encuestas electorales: aprendizajes del pasado proceso electoral¹

JAVIER ALAGÓN



La actividad demoscópica tiene altas dosis de riesgo. Son innumerables las posibles fuentes de error, especialmente porque dependemos de lo que la gente contesta a cuestionarios con flujos extraños, con escalas que les pueden parecer distantes y con preguntas que pueden ser inquietantes.

Sin embargo, a pesar de su compleja terminología, de lo desconcertante que puede parecer el pensamiento inductivo y de los desaciertos que llegan a ocurrir, el método estadístico seguirá siendo la principal forma de medición y entendimiento de formas de pensamiento y de comportamiento de poblaciones humanas. Y esto es así porque el método estadístico es, en realidad, parte del método científico.

El cuestionamiento sobre las encuestas en el pasado proceso electoral se refiere a que la mayoría de las encuestas publicadas por casas encuestadoras bien establecidas en nuestro país, presentaron un panorama en donde el PRI/PVEM ocupó una posición privilegiada a lo largo de muchos meses y semanas, guardando una clara ventaja sobre los demás partidos políticos.

De manera resumida, la ventaja establecida por la mayoría de las encuestas publicadas era de alrededor de 20 puntos, con una disminución a 15 puntos en las últimas mediciones, con claras excepciones.

Es importante reconocer que la mayoría de las encuestas publicadas sí reflejaron el 1-2-3-4 de los candidatos que se presentaron a elecciones, lo cual no deja de ser meritorio, dadas múltiples experiencias fallidas en otros países y en otros procesos electorales.

A pesar de lo anterior, quedó en el ambiente, tanto del círculo rojo como del ciudadano de la calle, que la fuerte diferencia referida, ampliamente publicada y difundida a lo largo de toda la campaña electoral, especialmente por medios masivos con gran alcance (televisoras y periódicos importantes), pudo haber influido de alguna manera en las intenciones de voto de los electores.

Nunca podremos establecer científicamente si esto ocurrió o no. Y si influyó, si lo hizo hacia arriba o hacia abajo en la diferencia entre candidatos punteros.

La discusión acerca de si la difusión de resultados de encuestas electorales influye en el comportamiento del electorado puede resultar muy interesante y aleccionadora. A nivel individual, es muy claro que sí influyen. Muchas personas han votado por opciones que incluso pueden ser ideológicamente distantes, con tal de

que determinado partido o personaje pueda o no pueda ganar.

La información moldea expectativas. Sin información de encuestas electorales, el sentido del voto muy seguramente sería distinto de los resultados finales en múltiples procesos. Pero la información tiene matices: puede ser objetiva y precisa, sin sesgos, con márgenes de error pequeños y con alto nivel de confiabilidad, o puede tener variantes de distintos colores y magnitudes, con sesgos de medición naturales o inducidos, con altos márgenes de error, con errores no muestrales considerables.

La difusión de la información con tantos y tantos matices, con tantas y tantas variantes puede confundir, engañar y enfadar al público. Y de ahí, la importancia de difundir encuestas realizadas con rigor estadístico e independencia.

¿Qué tanto influyó la publicación constante y sonante de múltiples encuestas electorales? No se ha podido, ni se podrá establecer de manera rigurosa. Pero seguramente, la expectativa de una diferencia muy abultada a favor del PRI/PVEM influyó en el ánimo y en la conducta de muchos votantes, aunque no sabemos si dicha influencia jugó a favor o en contra.

La distancia de 6.5 puntos porcentuales reflejada la misma noche del primero de julio sorprendió a tirios y troyanos, después de los 15 y 20 puntos de distancia a los que nos acostumbramos a lo largo de muchas semanas.

¹ Extracto de ponencia presentada en el IFE, en Seminario sobre Encuestas Electorales, noviembre 2012.



Los encuestadores cuestionados, colegas profesionales de la demoscopia, han esgrimido diversos argumentos en su defensa, desde que las encuestas no pueden ser usadas para predecir resultados electorales, que no son adivinos, que las encuestas reflejan las intenciones de voto sólo en el momento de la medición, hasta que las tasas de respuesta en México son muy bajas, que los sesgos de respuesta son especialmente complejos en nuestro país, que una proporción importante de mexicanos ocultan o distorsionan deliberadamente sus intenciones de voto, que las preferencias cambiaron de último momento en la última semana, justo cuando ya no se pueden publicar encuestas.

Otros argumentos de defensa se dirigen a los márgenes de error que –si se les considera– las diferencias ya no fueron tan grandes, que la asignación de indecisos es muy compleja, que el uso de modelos de asignación

de indecisos y de no respuesta es un proceso arriesgado, que se reportaron entonces las preferencias tal cual fueron expresadas por los entrevistados, que en otros países las cosas se han puesto peores, mucho peores.

Por supuesto que los argumentos anteriores pueden ser muy válidos y seguramente algunos o muchos de ellos influyeron para que a lo largo de varios meses la mayoría de las mediciones electorales reflejaran una realidad aparentemente muy simple.

Como estadístico no estoy satisfecho con el contraste mostrado por la mayoría de las encuestas electorales y el resultado final. La experiencia reciente en Estados Unidos es una demostración fehaciente de que, bien usada, la estadística no falla.

Después de muchas revisiones metodológicas y discusiones con colegas,

mi perspectiva estrictamente estadística es la siguiente:

1 La mayoría de las encuestas electorales de 2012 se quedaron cortas desde un punto de vista metodológico, principalmente en diseños de muestreo y levantamientos de campo. Los esquemas de muestreo convencionales utilizados en México desde hace varios lustros para encuestas electorales resultan insuficientes para la complejidad que vivimos en este país. Nuestro país (y muchos otros) representa un gran reto para que los encuestadores puedan obtener estimaciones sólidas, confiables y precisas de los parámetros de interés. Entre otras razones de esta complejidad se encuentran las siguientes:

- a. Localidades, municipios, regiones con altos niveles de peligro para encuestadores. Simplemente no se pueden hacer encuestas en



muchos lugares de este país. Tenemos muchos casos dolorosos de ataques registrados contra encuestadores en este sentido.

- b. Hay zonas pertrechadas a las que los encuestadores no tienen acceso, múltiples zonas residenciales cerradas, a las que es imposible acceder y por tanto, aunque hayan caído en muestra, no se pueden encuestar.
- c. Bajas tasas de respuesta por temor o por inseguridad en el manejo que se pueda hacer de la información.
- d. La geografía de nuestro extenso país hace especialmente difícil hacer levantamientos en localidades alejadas o muy remotas.

2 Existe documentación de sesgos de respuesta favorables al partido político que gobierna en algunos estados de la República Mexicana. En algunos estados, el sesgo de respuesta a favor del PRI fue hasta de 8 puntos. Esto significa que un grupo de personas encuestadas, que representan a una proporción no despreciable de la población mexicana, contestaba favorablemente a un partido político, sin tener la convicción de votar por dicho partido. Es decir, una parte (probablemente pequeña) de la explicación de la discrepancia referida entre encuestas y resultados finales pudo haber sido provocada por este sesgo. Como estadísticos, tenemos que aprender de estos sesgos.

3 La mayoría de las casas encuestadoras afirman que reportan los resultados tal cual son proyectados por métodos estadísticos, sin procesos de filtrado o de uso de modelos que en otras áreas y en otras latitudes han demostrado ser muy efectivos. En investigación de mercados, disciplina hermana de la opinión pública, el no hacer algún filtrado (por ejemplo, personas con intención dudosa de compra) es un camino directo al fracaso. El no utilizar modelos proyectivos empíricos no es la mejor estrategia. Es muy importante contar con modelos basados en experiencias que permitan ajustar las estadísticas derivadas de manera directa.

4 Los márgenes de error reportados, el típico de 3.1% al 95% de confianza para una muestra de tamaño 1,000, no corresponden a los diseños de muestreo utilizados en México. Estos márgenes corresponden a los de muestreo aleatorio simple, muy alejado de los muestreos que se hacen en México. Los márgenes de error reportados debieran ser mayores. Por tanto, las mediciones con encuestas electorales hechas con diseños convencionales, tienen mayor margen de error. ¿Cuánto? No se sabe, pero quizás hasta del doble.

5 Los diseños utilizados para *trackings* semanales, diarios o mensuales, deben estar basados en selecciones independientes y no en selecciones aleatorias iniciales que después se van rotando por conveniencia en campo. Esto podrá costar menos, pero implica menor calidad de información. Las selecciones independientes de muestras implican costos mayores, pero mucho mayor calidad de la información.

6 Por último, la presión a la que somos sometidos el gremio en términos de velocidad y costo, tiene, a su vez, un costo en la precisión. Es muy importante que en el futuro tomemos en cuenta que la calidad en encuestas electorales cuesta, y cuesta mucho.

Hemos tenido muchos aprendizajes en el gremio. Vislumbro un futuro prometedor basado en dichos aprendizajes, con muestras más sólidas, con mayor cobertura, con mayor empleo de filtrados y modelos, con mayor utilización de métodos de pronóstico. Simplemente, si hubiésemos aplicado los métodos convencionales de pronóstico para la mayoría de las series de tiempo reportadas por la mayoría de las casas encuestadoras, se hubiera obtenido una mejor estimación de la diferencia entre los dos candidatos punteros.

Algunos meses antes de las elecciones presidenciales pasadas, publiqué en la revista *DDT* de la AMAI, el artículo titulado “¿Cómo sobrevivir la tormenta? Una guía para leer encuestas en el México de 2012”. A manera de resumen y sin pretender argumentar todo aquello que se argumenta en dicho artículo, resumí en siete puntos lo que debemos tener en cuenta:

1. **Casa encuestadora con trayectoria, con reconocimiento en el gremio.**
2. **Encuesta pagada preferentemente por un medio de comunicación o alguna institución sin sesgos partidistas.**
3. **Con levantamientos cara a cara, ya sea en viviendas o por intercepción en la vía pública, de acuerdo a un esquema estratificado.**
4. **Con fichas metodológicas completas en donde se incluya el tipo de muestreo y el margen de error estimado de acuerdo al tipo de muestreo, no solamente el convencional de muestreo aleatorio simple.**
5. **Con tamaños de muestra preferentemente por encima de las 3,000 entrevistas en más de 300 secciones electorales.**
6. **Con selección aleatoria independiente de secciones electorales cada vez que se hace una medición.**
7. **Con distribución nacional de las secciones electorales para poder asegurar cobertura nacional, con muchas ciudades y localidades del país (de acuerdo a las selecciones aleatorias implicadas por el tipo de diseño muestral).**

Soy optimista. En la AMAI hemos revisado a profundidad el trabajo de nuestros colegas. Pensamos que puede hacerse un mejor trabajo en el futuro, pero nunca estaremos exentos de los riesgos inherentes en la demoscopia, método de ciencia y arte, empirismo e inspiración.



Javier Alagón es doctor en estadística por la Universidad de Oxford y matemático por la UNAM (Medalla Gabino Barreda). Su orientación científica por formación y por convicción es su principal filosofía en el fascinante mundo de la demoscopia.

