

Las encuestas presidenciales en México durante el Siglo XXI

Peña, Ricardo de la

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Empfohlene Zitierung / Suggested Citation:

Peña, R. d. I. (2015). Las encuestas presidenciales en México durante el Siglo XXI. *Revista Mexicana de Estudios Electorales*, 15, 78-114. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-459241>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier: <https://creativecommons.org/licenses/by/4.0/deed.de>

Terms of use:

This document is made available under a CC BY Licence (Attribution). For more information see: <https://creativecommons.org/licenses/by/4.0>

Las encuestas presidenciales en México durante el Siglo XXI.

Ricardo de la Peña¹

RESUMEN

Durante el presente siglo, ha habido tres ocasiones en que se han celebrado comicios para la elección de Presidente de la República en México, manteniendo una normalidad de periodos sexenales ininterrumpidos. En todas estas ocasiones, las encuestas han jugado un importante papel como orientadoras al electorado del formato de la contienda, pero su labor no ha estado exenta de críticas, no solo por la cantidad de mediciones publicitadas, sino por los niveles de inexactitud y por los sesgos percibidos en sus resultados por políticos, académicos, comentaristas, analistas y público interesado en general. Revisar la magnitud de estas inexactitudes y la relevancia de los sesgos existentes es una tarea pendiente, que puede ser atendida, aunque los motivos de los errores puedan seguir en suspenso, pues es sumamente difícil alcanzar, más allá del diagnóstico, un conocimiento cabal de los elementos intervinientes que pudieran haber propiciado las diferencias entre lo estimado mediante las encuestas en un momento determinado a una población dada de los resultados que se arrojen en las urnas por un grupo de votantes que enfrentan circunstancias distintas en un momento diferente y que no constituyen nunca, todo el universo de electores con derecho a votar.

Palabras clave: elecciones, presidencial, México, encuestas, exactitud, sesgo.

¹ Estudios de maestría en sociología en la Facultad de Ciencias Políticas y Sociales de la Universidad Nacional autónoma de México, Diplomado en Análisis Político en la Universidad Iberoamericana, Diplomado en Economía Aplicada en el Instituto Tecnológico Autónomo de México. Presidente Ejecutivo de *Indagaciones y Soluciones Avanzadas, S.C. (ISA)*. Correo electrónico: ricartur@gmail.com

ABSTRACT

79

During this century, there have been three occasions when elections were held for the election of President of the Republic in Mexico, maintaining a normal consecutive six-year periods. On all these occasions, surveys have played an important role in guiding the electorate format of the contest, but their work has not been without criticism, not only by the number of measurements publicized, but levels of inaccuracy and bias perceived at the results by politicians, academics, commentators, analysts and interested public. Check the extent of these inaccuracies and relevance of existing biases is a pending task, which can be addressed, although the reasons for the errors can still pending, it is extremely difficult to achieve, beyond diagnosis, a thorough knowledge of the intervening elements that may have led to differences between estimated by surveys in a particular to a given throw the results at the polls by a group of voters who face different circumstances at a different time and that all people are never far the universe of voters entitled to vote.

Key words: elections, presidential, Mexico, polls, accuracy, bias.

INTRODUCCIÓN

Durante el Siglo XXI México ha tenido tres elecciones de carácter presidencial, en las cuales no solamente se ha dado la alternancia en el Poder Ejecutivo Federal, sino que han sido contiendas reñidas, de difícil conocimiento adelantado de su resultado. En todas estas elecciones, las encuestas han jugado un papel importante no solamente por la cantidad de mediciones difundidas, notable y tendencialmente creciente, desde entonces, además por su condición de referente común para la discusión sobre el estado de la competencia en el círculo más atento (políticos, académicos, comentaristas, analistas) e inclusive, en el electorado en general, a lo que habría que sumar el hecho de que en la más reciente elección, las encuestas se convirtieron en un elemento central en la disputa formal del resultado.

Hacer un recuento de las encuestas publicadas y de sus estimaciones no es necesariamente, una tarea sencilla y nada impide que sea polémica e incluso equívoca. Son muchos los elementos a considerar para la inclusión, o no, de una medición y la determinación de cuáles de sus datos deben ser considerados

como los adecuados. Luego, para el tratamiento de esta información, en el caso de este artículo, se tomaron dos decisiones centrales: primera, homologar los datos de las distintas encuestas, excluyendo de los repartos los casos no definidos, lo que en muchas ocasiones, significó tomar el dato correspondiente al reporte generado por la propia encuestadora, aunque en otras, obligó a efectuar cálculos sobre los datos brutos aportados; y segunda, a tomar un único dato para cada encuestadora en cada elección: su última estimación pública, tomando los repartos producto de modelos de votante probable sobre preferencias efectivas, cuando se reportaron como tales, asumiendo que la encuestadora responsable optó por aproximarse a una estimación mediante la decantación de votantes entre los electores, desechando con ello, toda presunción de correspondencia entre preferencias efectivas y resultados por venir.

Igualmente, tuvo que definirse bajo qué principios o reglas incluir o excluir un estudio en particular. Así, se consideró únicamente la última encuesta publicada por cada firma, pero bajo tres condiciones adicionales: primera, que el dato correspondiente hubiera sido producto de una encuesta nacional con al menos un millar de casos observados; segunda, que el dato correspondiera a un ejercicio cuya toma de datos, en campo, hubiera concluido dentro de las cinco semanas inmediatas previas a la elección y no con anterioridad; y tercero, que no incluyera entrevistas por vía telefónica, debido a que de acuerdo con cánones internacionales, no resulta adecuado efectuar ejercicios por este medio en México, dada la relativamente reducida cobertura del sistema telefónico. Estas decisiones tratan de equilibrar las condiciones externas a la generación de las estimaciones, evitando cotejar datos que pudieran ser caducos, con escasa precisión o cuyo modo de recolección no sea homogéneo. La disposición de una ventana temporal definida reduce la posibilidad de que las diferencias entre estimaciones puedan atribuirse al momento de terminación del proceso de acopio de datos, reconociendo que el lapso entre una medición y la elección es una fuente primordial de diferencia de las encuestas con los resultados (Crespi, 1988: 183-184).

Este recuento no resulta original, pues es imposible reinventar la historia y conformar series nuevas de mediciones, antaño hechas, del dominio público. Tampoco lo es la toma de criterios que, a partir de ciertos principios lógicamente sustentables, dan lugar a una discriminación-selección unívoca. Luego,

lo novedoso de este artículo no será el recuento de mediciones, sino el tratamiento que de ellas se hace: ir más allá de la reflexión sobre los promedios observados en las encuestas publicadas, hacia la revisión crítica de las desviaciones y el sentido de estas desviaciones que se han presentado a lo largo del tiempo, en busca de sesgos sistemáticos y de la determinación, no solo de potenciales factores explicativos sino de mecánicas correctivas eventuales.

1. LAS ENCUESTAS PUBLICADAS.

Una primera observación, producto de este recuento es que si bien ha habido indudablemente un incremento del número de estudios registrados ante la autoridad administrativa electoral federal, durante los tiempos en que formalmente, se ha demandado la presentación de reportes (que no son necesariamente coincidentes con las duraciones de los procesos legales o de las campañas), ello fue producto de diversos fenómenos no necesariamente, vinculados con una elevación del número real de mediciones existentes, sino como producto de un aumento en la proporción de encuestas que fueron reportadas a la autoridad respecto a las realizadas. Lo que es más relevante: la mayor cantidad no significó mayor diversidad de proveedores de información, pues el volumen de casas encuestadoras que han difundido en cada proceso electoral federal se ha mantenido relativamente constante durante este siglo. Lo que ha variado es la frecuencia en que algunas encuestadoras realizan mediciones. Así, en 2012 por vez primera se contó con dos series de mediciones semanales y se difundió un ejercicio pionero de seguimiento diario de carácter público, que por sí solo tuvo cerca de un centenar de estimaciones, alrededor de la mitad del total reportado. Esta serie diaria, junto con las semanales, generaron tantas estimaciones diferentes como lo publicado seis años antes.

De hecho, las nuevas reglas para publicación (INE, 2014), de prevalecer sin cambios sustanciales para el próximo proceso electoral presidencial, bien podrían propiciar que el número de mediciones, a reportarse, descienda drásticamente, por varios motivos que es pertinente explicitar: primero, ante el vislumbrado equívoco de los programas de medición cotidiana, es factible que la disposición de mediciones sumamente frecuentes descienda o desaparezca, reduciendo el número de observaciones a disposición del público; pero, lo más importante, la complejidad que representa la obligatoriedad de dar a conocer

y constatar mediante elementos fiscalmente sólidos, los mecanismos de financiamiento dificultan la publicación de resultados de estudios por las vías controladas (que no incluyen el ciberespacio) no únicamente por supuestos falseamientos o encubrimientos, sino sobre todo por la complejidad misma del procedimiento y por la obligación de revelar montos que suelen, en muchas ocasiones, ser cubiertos con posterioridad o parcialmente por los pretendidos patrocinadores.

Siguiendo las reglas de selección impuestas, lógicamente sustentadas, pero cuestionables, como toda toma de decisión al respecto, tenemos colecciones homogéneas en cantidad de mediciones, presentadas al cierre de cada proceso electoral: catorce estudios en cada caso². Empero, el debate que se despertó en cada ocasión fue muy variado en contenidos, alcances y temáticas. En el año 2000, el debate sobre los motivos de los resultados, en busca del encuentro de una narrativa común que diera cuenta de lo ocurrido en la contienda, llevó al Instituto Federal Electoral, junto con otras asociaciones civiles, a promover un Taller en Sumiya, que reunió al gremio prácticamente en su totalidad. Sin embargo, las diferencias persistieron entre quienes apuntaban la viabilidad de detectar el ordenamiento adecuado, mediante el uso de las herramientas apropiadas y quienes afirmaban lo imprevisible mediante encuestas del resultado. A pesar de ello, hubo diversos avances que no pueden soslayarse, como el rechazo generalizado a los métodos de aproximación en vía pública y la definición de elementos sobre el diseño de muestras que irían alimentando lo que poco a poco ha ido conformando el canon metodológico de la investigación en el campo. Pero otros puntos no quedaron solventados entonces ni hasta ahora, tales como la pertinencia de ajustarse a los datos directamente obtenidos o procurar estimaciones con base en modelos de “votantes probables”, bien sean mediante procesos de filtrado de quienes se consideren como tales, bien a través de ponderaciones que otorguen probabilidades de concurrencia a sufragar a cada entrevistado.

² Las fuentes utilizadas para estos recuentos serían distintas en cada ocasión, según se ha ido avanzando en la conformación de concentrados públicos de encuestas. Para la elección presidencial del 2000, la mejor y la más completa fuente es la memoria del Taller de Sumiya (Taller de Sumiya, 2001). Tres años después lo fueron las memorias de un seminario organizado por la autoridad federal electoral (Instituto Federal Electoral, 2010); y para la elección presidencial de 2012 ya lo fue una fuente oficial (Instituto Federal Electoral, 2012).

El cuestionamiento sobre la que se consideraba relativa laxitud en la regulación de las encuestas electorales en el país, llevó a un debate público, del que inclusive, formó parte el legislativo federal, derivó en un claro mejoramiento y un fortalecimiento de las normas a aplicarse para posteriores contiendas electorales, forjando una relación normada y exigida entre autoridades electorales federales y organismos gremiales de la demoscopia para la definición de los criterios general de carácter científico para la publicación de encuestas. En esa ocasión, la mayoría de las encuestas publicadas daban una ventaja, corta en ocasiones, pero ventaja al fin y al cabo, a quien resultaría segundo lugar en la elección, invirtiendo el orden de los contendientes y errando en el ganador. Ello no fue genérico, pues al menos seis de las catorce mediciones seleccionadas, dieron un ordenamiento adecuado, con distancias entre los primeros lugares, no muy distintas a las registradas efectivamente.

CUADRO 1
ESTIMACIONES REPORTADAS POR LAS ENCUESTAS FINALES EN 2000

Encuestas finales 2000		Casos en muestra	ESTIMACIÓN REPORTADA			
Encuestadora	Patrocinador		VFQ	FLO	CCS	Otros
ARCOP	P.A.N.	1400	42.5	38.1	16.7	2.7
GEA-ISA	Autofinanciada	2400	40.5	38.2	18.0	3.3
Democracia y Watch	Particular	1542	41.0	36.0	20.0	3.0
GAUSSC	P.A.N.	1500	40.6	38.5	19.1	1.8
Demotecnia	Autofinanciada	2054	44.0	34.0	16.0	6.0
Alduncin y Asociados	Particular	2095	40.5	34.6	20.3	4.6
Grupo Reforma	Grupo Reforma	1545	39.0	42.0	16.0	3.0
AC Nielsen	Milenio Diario	2489	39.0	42.0	16.0	3.0
MERCAEI	Autofinanciada	1316	38.4	42.9	16.1	2.6
Consultores y MP	Autofinanciada	1800	38.0	41.0	19.0	2.0
CEO	Autofinanciada	2423	38.9	42.7	15.3	3.1
Zogby	Reuters	1330	40.7	43.6	14.5	1.2

Pearson	P.R.I.	1590	38.6	43.2	14.8	3.4
MUND	The Dallas Morning	1362	35.1	36.2	26.6	2.1
Promedio de encuestas	1911	39.8	39.5	17.7	3.0	
Resultado oficial	43.5	36.9	17.0	2.6		

Fuente: Cálculos del autor con base en la información de El papel de las encuestas en las elecciones federales. Memorias del Taller Sumiya 2000, IFE-Asociación Mexicana de Agencias de Investigación de Mercado y Opinión Pública-Colegio Nacional de Actuarios, México, 2001.

CUADRO 2

ESTIMACIONES REPORTADAS POR LAS ENCUESTAS FINALES EN 2006

Encuestas finales 2006		Casos en muestra	ESTIMACIÓN REPORTADA			
Encuestadora	Patrocinador		FCH	RMP	AMLO	Otros
GEA-ISA	Autofinanciada	1600	38.0	23.0	36.0	3.0
ARCOP	P.A.N.	1400	37.0	25.0	34.0	4.0
Grupo Reforma	Grupo Reforma	2100	34.3	25.3	36.4	4.0
Consultores y MP	Autofinanciada	1200	36.8	25.8	33.9	3.5
El Universal	El Universal	2000	34.0	26.0	36.0	4.0
CEO	No especificado	2000	33.5	25.3	35.8	5.4
Consulta Mitofsky	Televisa	2800	33.0	27.0	36.0	4.0
BGC Ulises Beltrán	Autofinanciada	1200	34.3	26.3	34.3	5.1
Parametría	Excélsior	1000	33.0	27.0	37.0	3.0
GAUSSC	No especificado	3250	34.6	28.8	33.7	2.9
Alduncin y Asociados	Particular	2046	35.0	29.0	32.0	4.0
Zogby	U. de Miami	1000	35.0	28.0	31.0	6.0
Demotecnia	Milenio Diario	2000	30.5	29.6	35.4	4.5
INDEMERC	O v a c i o n e s - ACIR	1500	32.3	28.2	33.5	6.0
Promedio de encuestas	1793	34.4	26.7	34.6	4.2	

Resultado oficial	36.9	23.0	36.3	3.8		
-------------------	------	------	------	-----	--	--

CUADRO 3

ESTIMACIONES REPORTADAS POR LAS ENCUESTAS FINALES EN 2012

85

Encuestas finales 2012		Casos en muestra	ESTIMACIÓN REPORTADA			
Encuestadora	Patrocinador		JVM	EPN	AMLO	GQT
MERCAEI	Autofinanciada	1200	27.2	38.5	31.0	3.3
Covarrubias y Asociados	SDP Noticias	1500	26.0	40.0	30.0	4.0
Demotecnia	UNO TV	1500	22.9	40.2	32.4	4.5
Grupo Reforma	Grupo Reforma	1616	24.0	41.0	31.0	4.0
Berumen y Asociados	O.U.E.	3480	22.4	41.7	34.0	1.9
IPSOS-BIMSA	Autofinanciada	1000	24.6	44.1	29.5	1.8
Consulta Mitofsky	Autofinanciada	1000	24.1	44.5	29.4	2.0
BGC Ulises Beltrán	Excélsior	1200	25.0	44.0	28.0	3.0
ARCOP	P.A.N.	1200	31.0	39.0	27.0	3.0
Parametría	O.E.M.	1000	23.6	43.9	28.7	3.8
Buendía & Laredo	El Universal	2000	24.4	45.0	27.9	2.7
Con Estadística	Grupo Fórmula	1150	24.7	44.4	26.7	4.2
GEA-ISA	Grupo Milenio	1144	22.4	46.9	28.5	2.2
INDEMERC	El Financiero	2000	22.8	47.2	27.1	2.9
Promedio de encuestas		1499	24.7	42.9	29.4	3.1
Resultado oficial			26.1	39.2	32.4	2.3

Fuente: elaboración propia con base en la información del: *Séptimo Informe que presenta la Secretaría Ejecutiva al Consejo General del Instituto Federal Electoral respecto del cumplimiento del Acuerdo CG411/2011*, IFE, México, 26/07/2012

Luego, si el debate público y el análisis, supuestamente experto, de los resultados de los ejercicios, no coincide con la evidencia empírica inicial, cabe preguntarse cómo medir la exactitud de las encuestas electorales publicadas en México

durante el siglo que corre. ¿Existe un único criterio universalmente válido? La respuesta es obvia y ya la externó el especialista Warren Mitofsky (1998): “Después de más de cincuenta años de encuestas electorales (...) no ha sido adoptada por la comunidad de encuestadores una métrica estándar para medir la exactitud de las encuestas.” Añadiríamos para precisar aún más: no solo no se ha consensuado entre expertos; más bien no ha sido hallada una métrica estándar para medir este fenómeno. Busquemos ahora avanzar en este terreno.

2. EL ERROR COMO DIFERENCIA³

Son muy diversos los posibles estimadores para medir la proximidad entre un grupo de datos y otro. En el caso del cotejo entre estimadores, producto de encuestas y resultados de una elección, son muchas las propuestas presentadas, aunque pocas las realmente utilizadas. Idealmente, un estimador de la exactitud de una encuesta o de la diferencia entre lo estimado y el parámetro, debiera cumplir algún conjunto de criterios convenidos *a priori*, con un mínimo de sentido común. A manera de ilustración, sin que exista necesariamente un consenso al respecto, se someten a examen las siguientes condiciones:

Primero: debiera posibilitar la comparabilidad entre mediciones y de éstas con un valor teórico esperado, bien para un conjunto de encuestas en una elección, para diversos ejercicios de una misma firma o para una colección de encuestas en elecciones diversas, permitiendo un análisis agregado, a la vez que un tratamiento particularizado que determine las relaciones entre distintas mediciones.

Segundo: debiera generarse por un procedimiento homogéneo para todos y cada uno de los contendientes, cualquiera que sea su número, siendo aplicable a resultados tanto en sistemas bipartidistas como multipartidistas; y considerar además, la totalidad de las opciones de respuesta predeterminadas y dadas a conocer al encuestado de respuesta al reactivo, sobre preferencias, usado para fines de estimación, así como aquéllas que se hubieran dado de manera espontánea, adicionales a las opciones acotadas por el entrevistador.

3 Este apartado, en la primera parte, de índole expositiva, retoma en lo fundamental, aunque en un orden distinto por razones prácticas, el contenido de un ensayo del autor sobre cómo medir la exactitud de las encuestas electorales, en preparación y con la pretensión de ser publicado como texto de carácter técnico.

Las propuestas más conocidas y usadas datan de hace más de medio siglo y en todos los casos reclaman la disposición de dos grupos de datos: las proporciones por contendiente, observadas en los estudios y las proporciones oficiales, obtenidas por los mismos contendientes. Es muy importante recordar al respecto que, en estadística, precisión y exactitud no son conceptos equivalentes, aunque estén relacionados. La precisión es una medida de la dispersión del conjunto de valores obtenibles en repetidas mediciones de una magnitud determinada: a menor dispersión, mayor precisión. Luego, refiere a la magnitud escalar con la que se mide la proximidad de los resultados de una muestra para reproducir los resultados que se obtendrían de un recuento completo llevado a cabo usando las mismas técnicas. El diseño de una buena encuesta probabilística busca la generación de estimadores con la menor dispersión posible, dadas las condiciones concretas para su realización.

La exactitud, en cambio, es una medida de la proximidad entre los valores medidos y los reales, aquellos resultantes, si se hubiera realizado un censo bajo los mismos procedimientos que la encuesta, y cuando la diferencia es distinta de cero, mide el sesgo de una estimación determinada: a menor sesgo, mayor exactitud. Luego, refiere a la magnitud escalar con la que se mide la proximidad del estimador producto de una muestra respecto del parámetro objetivo. El diseño de la encuesta puede privilegiar la adopción de estimadores teóricamente a-sesgados o bien, utilizar estimadores sesgados, pero con alta precisión, según lo que se considere óptimo en el diseño.

Fue la experiencia de desencuentro entre las estimaciones por encuesta y los resultados oficiales de la elección presidencial de 1948 en Estados Unidos la que llevó a una revisión profunda de los ejercicios, que incluyó un análisis, a cargo de Frederick Mosteller (1949), sobre los posibles métodos para medir el error de las estimaciones. En ese texto clásico, el autor lista varios métodos que considera factibles y eventualmente pertinentes.

Es preciso mencionar que un problema que Mosteller no atiende, en su momento, debido a que aún no había surgido como tal, y que ha sido fuente de divergencias y discusiones en las evaluaciones posteriores, respecto al error en encuestas, es el relativo a qué datos considerar para fines de análisis. De nuevo, es un debate en torno a la supuesta inadecuación entre encuestas y resultados, ocurrido medio siglo más tarde, en las elecciones presidenciales de Estados

Unidos en 1996, lo que coloca en la palestra la discusión sobre el problema de los casos no definidos. Al igual que en el caso de la estimación de la exactitud de las encuestas, no existe hasta ahora, un consenso respecto al tratamiento pertinente de los casos no definidos, aunque sí han surgido propuestas de estimadores que buscan expresamente, resolver de manera práctica este asunto.

Los reportes de resultados de las encuestas electorales pueden referir a diversas salidas de información: una primera dicotomía es la existente entre proporciones observadas, incluyendo como un componente diferenciado los casos no definidos por algún contendiente, contra las proporciones efectivas, que excluyen los casos no definidos o suponen su asignación. Otra dicotomía es entre los datos directamente obtenidos, con las ponderaciones acordes al diseño adoptado, y los datos ajustados conforme a la probabilidad estimada de sufragio, producto de los llamados modelos de “votantes probables” u otras opciones para el cálculo de proporciones.

En el caso de Estados Unidos, las proporciones de casos no definidos suelen situarse en niveles relativamente reducidos, por lo que las diferencias en los cálculos de error entre los datos observados y los resultantes de la exclusión o asignación proporcional de no definidos, no es tan relevante. Ello tiene como trasfondo una visión de los profesionales estadounidenses, para los cuales el tratamiento pulcro de los datos debiera privilegiar el reporte de lo observado tal cual fue recuperado, sin manipulación alguna, pero que a la vez concibe los niveles de no definición en la pregunta electoral, como reflejo de problemas en el diseño de instrumentos o de carencias del debido rigor en los procedimientos operativos, lo que sería al menos cuestionable; de hecho, suele considerarse pertinente que el investigador deba realizar esfuerzos para lograr niveles de respuesta elevados.

Empero, en otras democracias, la proporción de casos no definidos en una encuesta, suele ubicarse en niveles más elevados y considerarse reflejo de una actitud real de indefinición del elector y no como un problema de técnicas de acopio de datos, por lo que no efectuar una exclusión o una asignación proporcional, lleva a estimaciones poco diáfanas de los niveles de error registrados. Es por ello que antes de un tratamiento para el cálculo de los errores en las estimaciones, suele efectuarse una homologación de los datos, aplicando procedimientos de cálculo que excluyen o asignan los casos no definidos, mediante

algún procedimiento uniforme (o desproporcional si así lo determina la experiencia y creatividad del analista).

89

De los diversos métodos para el cálculo del error de las encuestas, resumidos por Mosteller, siete de ellos resultan de interés: recuperar, por basarse en la relación entre las proporciones de voto estimadas y observadas (un octavo refiere a la relación entre estimación y reparto real de votos electorales, lo que es propio solamente de un sistema de elección indirecta, como el estadounidense). La mayoría de ellos no dan respuesta alguna sobre el posible tratamiento de los casos no definidos. Mosteller pareciera asumir que se debieran tomar los datos tal cual se reportan, sin recalcular alguno, para excluir la proporción no definida en la muestra. En análisis posteriores, sin embargo, es usual que se recurra a procesos de asignación o exclusión, no siempre idénticos.

Hagamos un recuento de los métodos propuestos por Mosteller, proponiendo una notación que los precise y señalando sus alcances y limitaciones, siguiendo al autor y a otros analistas posteriores, particularmente Mitofsky. El recorrido lo realizaremos no siguiendo el orden originalmente dado en el texto por Mosteller, sino en razón del número de componentes considerados para la estimación del error, de menos a más, para clarificar la exposición. Además, dada su relevancia como aspecto generador de ambigüedad y divergencias entre los cálculos del error para idénticos conjuntos de encuestas, se recobrará la notación relativa a las dos opciones fundamentales de estimación: con los datos directamente obtenidos, incluyendo un segmento de casos con una intención de voto no definida; y con los datos relativos exclusivamente al segmento de casos definidos por algún contendiente.

Uno sería la diferencia en puntos porcentuales entre la proporción prevista para el ganador y la proporción oficial obtenida respecto al total de votos emitidos (M1). Este método resulta el más simple posible y ocupa solamente dos datos para fines de estimación del error, lo que lo hace intuitivo para cualquiera. Empero, como Mitofsky advierte, este método resulta artificial, sobre todo cuando el líder cuenta con menos de la mitad de los votos, por existir terceros contendientes o al incluirse al segmento de indefinidos.

Otro, la diferencia en puntos porcentuales entre las proporciones predichas y reales de votos recibidos por los dos mayores contendientes (M2). Este método,

sencillo de calcular y de fácil entendimiento, cuenta con la clara ventaja de ser el único de los propuestos por Mosteller que no se afecta por la inclusión o no, de los casos no definidos. Eso se logra, eliminando del cálculo las proporciones correspondientes a opciones menores, regularmente reportadas por separado. Mosteller privilegia los dos primeros métodos sobre otros, debido a problemas que surgen al aplicar las opciones restantes. Pero ello es consecuencia más que de una regla generalizable a todo sistema electoral, del carácter propiamente bipartidista de la realidad que atiende este autor: las elecciones en Estados Unidos.

Una más, la diferencia absoluta entre lo previsto y lo real para los dos mayores contendientes, respecto al total de votos emitidos (M5). Lo primero a destacar es que, en contiendas bipartidistas, M2 y M5 resultan ser equivalentes, si se asume la inexistencia o la asignación proporcional de los casos indefinidos. M5 corresponde, por demás, con lo regularmente resaltado en los reportes mediáticos: el margen de ventaja del líder. Por lo anterior, ha sido usado regularmente en los análisis sobre el tema, incluso por el ámbito académico estadounidense y, dado lo común de analistas mexicanos de seguir lo hecho por sus vecinos, no es extraño encontrarlo como el indicador usado y considerado pertinente, al margen de que elimina a una parte muy significativa de la composición del voto realizado. Sin embargo, al considerar las proporciones para los dos primeros lugares respecto al total de todos los componentes, pierde la virtud de arrojar un estimador que no se afecte por el problema de los casos no definidos.

Otro, la desviación media en puntos porcentuales entre lo previsto y lo real para todos los contendientes, sin tener en cuenta el signo de la diferencia (M3). Este tercer método, si bien toma en cuenta todos los componentes para generar un único estimador correspondiente a lo intuitivo y sencillo de calcular, no permite una comparación diáfana cuando el número de partidos es variable e impide reconocer el sesgo en la medición. Asimismo, el grado de error no se encuentra relacionado con las proporciones de voto obtenidos, sino que asume un valor absoluto, por lo que puede derivar en cálculos reducidos de error, al incluir a diversos componentes menores que poco contribuyen al voto y cuya estimación suele presentar divergencias absolutas pequeñas, por lo que promediarlos con los componentes mayores propicia una estimación inconveniente de la magnitud del error registrado.

Alternativamente, se ha adoptado M3 como método para el cálculo del error en las encuestas preelectorales, asignando los indefinidos según las proporciones para los definidos, evitando un incremento artificial del error en los estudios en que no se realiza el ejercicio de asignación y permitiendo un tratamiento homogéneo de los datos relativos a las diversas encuestas públicas. Empero, no suele incluirse el error correspondiente al remanente de “otros candidatos”, sino solamente de aquellos contendientes que son reportados por separado por las encuestadoras en cada elección.

La solución de definir cada ocasión el número de contendientes a considerar o asumir un umbral determinado para su inclusión en el cálculo, y luego promediar los errores observados, además de ser un mecanismo arbitrario, mantendría problemas de homogeneidad, bien en la cantidad de contendientes considerados, bien en la proporción del voto incluida en el cálculo, bien en ambos aspectos.

No deja de ser paradójico y relevante que las razones que sustentan la decisión de Mosteller de privilegiar para el análisis, métodos determinados para el cálculo de las desviaciones de las encuestas, respecto a los resultados, sean también los motivos para el empleo de métodos alternativos en situaciones distintas. El objetivo primordial de Mosteller, que a su parecer debiera guiar todo análisis, es la construcción de una base de datos homogénea. En el caso del sistema estadounidense, ello se logra cuando se consideran solamente dos componentes, excluyendo o colapsando, en lo posible, terceras opciones que regularmente, no están presentes o cuyo peso electoral es absolutamente marginal.

En sistemas multipartidarios no hay razones de peso para privilegiar métodos de agrupamiento o de reducción de componentes, únicamente a los dos mayores. De hecho, ni el margen de victoria, medido respecto a los votos por los dos partidos mayores o respecto al total de votos emitidos, ni mucho menos el solo voto por el ganador resultan procedimientos completos y correctos para calcular los niveles de desviación entre lo estimado y el resultado. En muchos análisis para sistemas multipartidarios se ha tendido entonces a privilegiar el empleo de M3, como el estimador más adecuado y completo para determinar la desviación de las encuestas respecto a los resultados, aún y cuando suela acompañarse de M5, reconociendo que es la diferencia entre los dos primeros

contendientes, la que suele privilegiarse por medios y público atento, como dato relevante producto de una encuesta, aunque ello sea equívoco.

Y es equívoco al menos por dos razones: no es un estimador directamente producido por la encuesta, sino el resultado de un cotejo de dos estimadores primarios (la preferencia por cada uno de los dos partidos mayores) y por ende, sujeto a un doble error, dada la desviación esperada para cada componente que se compara; y es parcial, al reducir el fenómeno de concordancia entre encuestas y resultado a la distancia medida entre dos componentes, excluyendo el peso observado en otros componentes, que si bien son menores, pudieron haber sido correcta o incorrectamente medidos, al margen de la corrección del margen de victoria estimado.

Está luego la media de las desviaciones de la razón entre la proporción prevista y la real, para todos los candidatos (M4). Con mayor complejidad para su cálculo, este cuarto método presenta el problema inverso a M3: tiende a darle un peso muy elevado a las desviaciones en componentes menores, dado que pequeñas diferencias tienen un alto peso al compararse con la proporción reducida de voto logrado por dichos componentes.

Además, se tendría la diferencia máxima observada entre lo previsto y lo real para cualquiera de los contendientes (M6). Este sexto método presenta problemas técnicos y puede propiciar confusión como estimador del rendimiento de los estudios, según advierte el propio Mosteller. Adicionalmente, podemos señalar que presenta serias limitaciones, no tanto en la comparabilidad entre estudios individuales, pero sí para toda pretensión de tratamiento agregado, pues no resulta adecuado para un cálculo de medias y el cálculo de máximos comunes ubica el error conjunto en un punto arbitrario.

Finalmente, se tendría un método más complejo, correspondiente a la prueba estadística chi-cuadrado, útil para evaluar la congruencia entre la distribución estimada y la distribución oficial de los votos (M7). El caso de este séptimo método, además de ser desestimado por Mosteller dada su complejidad y poca claridad para el público, enfrenta serias limitaciones para el tratamiento agregado de estimaciones diversas. Parte del problema, se resuelve recurriendo a cálculos a partir de las frecuencias de casos y no de proporciones, pero ello hace más compleja su estimación. Por ello, la complejidad y los problemas técnicos,

propios de un método como la prueba chi cuadrado o de opciones mejoradas de la misma, hace poco adecuado considerarlo como la opción pertinente.

93

Recientemente, se presentó la propuesta de un nuevo estimador del error en las encuestas, respecto al resultado (Martin *et. al.*, 2005, *passim*) que busca resolver el problema de tratamiento de los casos no definidos, mediante la eliminación de todo efecto del mismo en el cálculo. Empero, esta propuesta continúa siendo una medición adecuada solamente para sistemas bipartidistas, no para multipartidista; e incluso, uno con características muy peculiares, que se encuentran en muy pocos sistemas: la constancia temporal de la condición de dos partidos específicos como las opciones efectivamente competitivas, pues asume que existen dos partidos que son siempre los primeros lugares en una elección y, presumiblemente, en toda encuesta y, en todo caso, asigna una condición de tercer contendiente a cualquier otro participante en una elección

Esta propuesta parte del reconocimiento de que, si bien la diferencia entre las proporciones estimadas, considerando o excluyendo casos indefinidos, no son iguales, sí son iguales las razones entre las mismas proporciones. Asumiendo lo anterior, se construye un estimador del error (A) que toma el logaritmo natural del cociente de la razón entre las proporciones estimadas, y la razón de las proporciones reales para dos partidos. Con ello, los autores de esta propuesta eliminan el problema derivado de la adopción de cualquier método de asignación o supresión del segmento indefinido, pues carece de relevancia para el cálculo; permiten la detección simultánea de la magnitud y el sentido del error en la estimación, pues si su valor es cero, supone la coincidencia en el peso relativo entre los dos contendientes considerados, si el valor es positivo, corresponderá a una sobrestimación del partido cuyo error se calcula y si es negativo, supondrá una subestimación. Además, disponen de un estimador de la varianza y, por ende, del error estándar de la medición, por lo que es factible cotejar si la magnitud del error es estadísticamente significativa.

Sin embargo, cabe repetir, es una propuesta solamente útil en sistemas de dos partidos permanentes, lo propio del sistema estadounidense pero no necesariamente del resto del mundo. Como solución propuesta para la medición del error relacionado con terceros contendientes, se plantea la disposición de un cálculo suplementario e independiente del error, pero ello no responde a la inquietud, pertinente y propia de los análisis para sistemas multipartidarios, de

disponer de un estimador agregado del error de las mediciones que den cuenta del fenómeno, considerando todos los componentes, o al menos los relevantes, que no forzosamente son dos de los competidores.

Así, la revisión de los distintos medidores disponibles nos da una colección de críticas, pero parece no resolernos el problema. Es obvio que del conjunto de estimadores de exactitud varios podemos dejarlos de lado desde un principio: M1 y M6 por ser excesivamente parciales, al tomar un único componente; al igual A, complejo y claramente bipartidista; y M7, por ser de difícil interpretación para los legos y que fuera descartado inclusive por el propio autor de la mayoría de las propuestas. ¿Qué pasa con M2 y M5, tan socorridos en la literatura y adoptados para casos muy diversos? Su condición de contabilizar solamente los dos primeros lugares en una contienda, los hace claramente impropios para sistemas multipartidarios (aunque los seguidores de las escuelas y las costumbres estadounidenses no reparan en ello). Queda luego M3 y M4, uno fundado en una contabilidad directa (sumas y restas) y el otro buscando promediando las razones. Por su mayor proximidad a lo intuitivo, su sencillez y claridad, del conjunto de estimadores posibles de la (in)exactitud de las encuestas electorales, respecto del resultado, lo pertinente pareciera ser quedarse con M3 como medidor privilegiado, dentro de lo disponible hoy en día en la literatura.

Este tema no deja de ser un problema central en la valoración del trabajo demoscópico, pues, aunque es totalmente cierto que al evaluar la calidad de un estudio por encuesta lo pertinente es observar lo riguroso de los procedimientos utilizados, para el grueso del público interesado, sin embargo, lo relevante para evaluar la calidad de las encuestas electorales es su proximidad con los resultados, entendiendo las estimaciones por muestreo bajo un concepto radicalmente diferente al formal: el público interesado entiende los resultados de las encuestas, en el terreno electoral, como aproximaciones a eventos futuros, a diferencia de lo que pueden hacer los encuestadores; estimar eventos pasados o vigentes en el momento de la entrevista.

El público interesado busca luego disponer de un medidor del éxito, logrado por las mediciones. Y hablar de un medidor implica asumir que lo que se demanda es la búsqueda de un estimador agregado de la exactitud que dé cuenta

de este fenómeno, en un único número. Y en nuestro caso, optamos por tomar M3 como este medidor, por las múltiples razones antes esgrimidas.

95) ¿Qué pasa si aplicamos esta vara para medir los resultados de las colecciones de estudios previos a elecciones en México, durante el presente siglo? Lo primero que encontramos es que los niveles de desviación o de inexactitud, si se quiere, que se han venido observando en las últimas tres elecciones presidenciales, son relativamente constantes; 1.9 para el promedio de las encuestas en 2000, con un error medio medido por M3 de 2.6 en cada una; 2.1 de desviación acumulada en los componentes, como media en 2006, pero con un error medio por encuesta, nuevamente calculado por M3, de 2.3 puntos; y un error para el promedio de las encuestas de 2.2 en 2012, que cuando se estima como error, conforme a M3 en cada una de ellas, alcanza el nivel de 2.6 puntos, el mismo que el observado en 2000.

CUADRO 3

ERROR CONTABILIZADO DE LAS ENCUESTAS FINALES EN 2000

Encuestas finales 2000	Casos en muestra	ERROR CONTABILIZADO				M3
		VFQ	FLO	CCS	Otros	
Encuestadora						
ARCOP	1400	-1.0	1.2	-0.3	0.1	0.65
GEA-ISA	2400	-3.0	1.3	1.0	0.7	1.50
Democracy Watch	1542	-2.5	-0.9	3.0	0.4	1.70
GAUSSC	1500	-2.9	1.6	2.1	-0.8	1.85
Demotecnia	2054	0.5	-2.9	-1.0	3.4	1.95
Alduncin y Asociados	2095	-3.0	-2.3	3.3	2.0	2.65
Grupo Reforma	1545	-4.5	5.1	-1.0	0.4	2.75
AC Nielsen	2489	-4.5	5.1	-1.0	0.4	2.75
MERCAEI	1316	-5.1	6.0	-0.9	0.0	3.00
Consultores y MP	1800	-5.5	4.1	2.0	-0.6	3.05
CEO	2423	-4.6	5.8	-1.7	0.5	3.15
Zogby	1330	-2.8	6.7	-2.5	-1.4	3.35
Pearson	1590	-4.9	6.3	-2.2	0.8	3.55
MUND	1362	-8.4	-0.7	9.6	-0.5	4.80

Promedio de encuestas	1911	-3.7	2.6	0.7	0.4	1.86
Resultado oficial		3.7	3.4	2.4	0.9	2.62

Fuente: elaboración propia con base en datos de *El papel de las encuestas en las elecciones federales. Memorias del Taller Sumiya 2000.*

CUADRO 4
ERROR CONTABILIZADO DE LAS ENCUESTAS FINALES EN 2006

Encuestas finales 2006	Casos en muestra	ERROR CONTABILIZADO					
		FCH	RMP	AMLO	Otros	M3	
Encuestadora							
GEA-ISA	1600	1.1	0.0	-0.3	-0.8	0.55	
ARCOP	1400	0.1	2.0	-2.3	0.2	1.15	
Grupo Reforma	2100	-2.6	2.3	0.1	0.2	1.31	
Consultores y MP	1200	-0.1	2.8	-2.4	-0.3	1.40	
El Universal	2000	-2.9	3.0	-0.3	0.2	1.60	
CEO	2000	-3.4	2.3	-0.5	1.6	1.95	
Consulta Mitofsky	2800	-3.9	4.0	-0.3	0.2	2.10	
BGC Ulises Beltrán	1200	-2.6	3.3	-2.0	1.3	2.30	
Parametría	1000	-3.9	4.0	0.7	-0.8	2.35	
GAUSSC	3250	-2.3	5.8	-2.6	-0.9	2.90	
Alduncin y Asociados	2046	-1.9	6.0	-4.3	0.2	3.10	
Zogby	1000	-1.9	5.0	-5.3	2.2	3.60	
Demotecnia	2000	-6.4	6.6	-0.9	0.7	3.65	
INDEMERC	1500	-4.6	5.2	-2.8	2.2	3.70	
Promedio de encuestas	1793	-2.5	3.7	-1.7	0.4	2.09	
Resultado oficial		2.7	3.7	1.8	0.8	2.26	

Fuente: elaboración propia con base en datos de *Memorias del seminario: encuestas y elecciones 2006*, Instituto Federal Electoral-Asociación Mexicana de Agencias de Investigación de Mercado y Opinión Pública-Asociación Mundial de Investigadores de la Opinión Pública-Consejo de Investigadores de la Opinión Pública, México, 2010.

CUADRO 5

ERROR CONTABILIZADO DE LAS ENCUESTAS FINALES EN 2012

97

Encuestas finales 2012	Casos en muestra	ERROR CONTABILIZADO				M3
		JVM	EPN	AMLO	GQT	
Encuestadora						
MERCAEI	1200	1.1	-0.7	-1.4	1.0	1.05
Covarrubias y Asociados	1500	-0.1	0.8	-2.4	1.7	1.25
Demotecnia	1500	-3.2	1.0	0.0	2.2	1.60
Grupo Reforma	1616	-2.1	1.8	-1.4	1.7	1.75
Berumen y Asociados	3480	-3.7	2.5	1.6	-0.4	2.05
IPSOS-BIMSA	1000	-1.5	4.9	-2.9	-0.5	2.45
Consulta Mitofsky	1000	-2.0	5.3	-3.0	-0.3	2.65
BGC Ulises Beltrán	1200	-1.1	4.8	-4.4	0.7	2.75
ARCOP	1200	4.9	-0.2	-5.4	0.7	2.80
Parametría	1000	-2.5	4.7	-3.7	1.5	3.10
Buendía & Laredo	2000	-1.7	5.8	-4.5	0.4	3.10
Con Estadística	1150	-1.4	5.2	-5.7	1.9	3.55
GEA-ISA	1144	-3.7	7.7	-3.9	-0.1	3.85
INDEMERC	2000	-3.3	8.0	-5.3	0.6	4.30
Promedio de encuestas	1499	-1.5	3.7	-3.0	0.8	2.24
Resultado oficial	2.3	3.8	3.3	1.0	2.59	

Fuente: elaboración propia con base en datos del: Séptimo Informe que presenta la Secretaría Ejecutiva al Consejo General del Instituto Federal Electoral, respecto del cumplimiento del Acuerdo CG411/2011, Instituto Federal Electoral, México, 26 de julio de 2012.

¿Qué pasa si intentamos una agregación de las (in)exactitudes observadas en una única tabla que dé cuenta de los resultados obtenidos a lo largo del siglo, por las diversas firmas encuestadoras? Lo primero que observamos es que el cálculo está viciado por un fenómeno relevante: mientras cuatro casas han aportado estimaciones finales en todos los casos, algunas lo han hecho solo en una ocasión y las más en dos elecciones. Esto produce una disparidad en el riesgo que han enfrentado las diversas firmas y coloca en los extremos, casos con una única medición aportada, aunque ésta no sea la mejor (que corresponde

a la encuesta GEA-ISA de 2006), aunque sí sea la peor observación la situada al final de la tabla.

CUADRO 6

INEXACTITUD PROMEDIO OBSERVADA ENTRE LAS ESTIMACIONES POR ENCUESTA Y LOS RESULTADOS OFICIALES DE LAS ELECCIONES PRESIDENCIALES DE MÉXICO CONFORME AL MEDIDOR M₃ (2000-2012)

ORDEN	ENCUESTADORA	CASOS	ELECCIÓN PRESIDENCIAL			M3
			2000	2006	2012	
1	Covarrubias y Asociados	1			1.25	1.25
2	ARCOP	3	0.65	1.15	2.80	1.53
3	El Universal	1		1.60		1.60
4	Democracy Watch	1	1.70			1.70
5	Grupo Reforma	3	2.75	1.31	1.75	1.94
6	GEA-ISA	3	1.50	0.55	3.85	1.97
7	MERCAEI	2	3.00		1.05	2.03
8	Berumen y Asociados	1			2.05	2.05
9	Consultores y MP	2	3.05	1.40		2.23
10	GAUSSC	2	1.80	2.90		2.35
11	Consulta Mitofsky	2		2.10	2.65	2.38
12	Demotecnia	3	1.95	3.65	1.60	2.40
13	IPSOS-BIMSA	1			2.45	2.45
14	BGC Ulises Beltrán	2		2.30	2.75	2.53
15	CEO (U de G)	2	3.15	1.95		2.55
16	Parametría	2		2.35	3.10	2.73
17	AC Nielsen	1	2.75			2.75
18	Alduncin y Asociados	2	2.65	3.10		2.88
19	Buendía & Laredo	1			3.10	3.10
20	Zogby	2	3.35	3.60		3.48
21	Pearson	1	3.55			3.55

22	Con Estadística	1			3.55	3.55
23	INDEMERC	2		3.70	4.30	4.00
24	MUND	1	4.80			4.80
PROMEDIO			2.62	2.26	2.59	2.49

Fuente: cálculos del autor con base en los datos compilados en los cuadros 3, 4 y 5.

3. EL ERROR COMO NORMALIDAD⁴

Como vimos antes, los estadísticos para medir la exactitud (o si se prefiere el error), en las encuestas suelen aportar estimadores de la distancia, sin reparar en su relación con la desviación esperada en las estimaciones. Generar un estimador que sí lo haga pudiera lograrse mediante la simple introducción en el algoritmo de un procedimiento de normalización convencional. Esto se consigue generando un estimador del error que dependa, no de la magnitud absoluta de la diferencia entre lo estimado y lo real o de su cociente, sino de la magnitud relativa de esta diferencia, con el error esperado, tomando como medidor del error esperado, la desviación estándar de las proporciones estimadas, dado un tamaño de muestra útil para su estimación.

Estrictamente hablando, este estimador de la dispersión de los datos solo es pertinente cuando se trata de un muestreo aleatorio simple. Sin embargo, dado que éste es el método, con mínimas restricciones en el proceso de selección, suele ser utilizado como punto de referencia para evaluar la eficiencia de otros diseños por la simplicidad de su cómputo, al costo de ignorar que regularmente, no fue el diseño usado (lo que se debiera declarar normalmente, de manera explícita, aunque no se hace), e inclusive para estimar un error esperado de manera genérica en un experimento. Este error esperado, es básico para el cálculo del margen de error para un estimador de proporción, que no es otra cosa que el error esperado con una probabilidad de ocurrencia dada. El error esperado disminuye a medida que aumenta el tamaño de la muestra disponible. Igualmente, es menor a medida que la proporción de ocurrencia de un evento se aleja del inverso del número de resultados posibles. Empero, estos cambios no son lineales, sino parabólicos, con diferente escala, según el tamaño de la muestra de la que se derivan, como es de sobra conocido y referido. A partir del

⁴ Nuevamente este apartado, en su primera parte, retoma en lo fundamental, desarrollos propios y propuestas iniciales del autor sobre cómo medir la exactitud de las encuestas electorales, que formarían parte de un trabajo en preparación con la pretensión de ser publicado como texto de carácter técnico.

cálculo del margen de error, puede establecerse un intervalo dentro del cual es esperable se halle el parámetro con cierta probabilidad predeterminada. Cabe recordar al respecto que el nivel de confianza y la amplitud del intervalo, varían conjuntamente, de forma que un intervalo más amplio y menos preciso, tendrá más posibilidades de acierto, mientras que un intervalo menor, que ofrece una estimación más precisa, aumenta las posibilidades de error. (100

Si los contendientes son solo dos, el error esperado con un muestreo aleatorio simple es el máximo posible. Con base en este margen de error máximo, suele reportarse un margen de error genérico para una encuesta. El error máximo posible, correspondiente a eventos dicotómicos cuando es igualmente probable su ocurrencia o no, se asume también como el error medio esperable para un conjunto de estimadores cuya suma sea unitaria, como son las proporciones para los diversos contendientes en una elección, lo que es, por demás, claramente incorrecto.

Así, de manera errónea, se concluye que el margen de error medio máximo esperado para un grupo de estimadores de proporciones, cuya suma sea igual a la unidad, será equivalente al margen de error máximo de una estimación cualquiera, solo cuando se trate de eventos dicotómicos, pues de lo contrario, como en el caso de contiendas multipartidarias, será menor; y crecientemente menor a medida que aumente el número de contendientes. Disponiendo de un margen de error tolerado a un nivel de confianza dado, para cada estimación particular, producto de una encuesta, pudiera verse si su diferencia con el resultado estuvo dentro o fuera de lo esperado. Esto supone ver cómo cumple una condición dada, fácilmente determinable. Puede también calcularse la dispersión de las proporciones estimadas para un componente en un conjunto de estimaciones dado y de aquí calcularse el margen de error para un conjunto de estimaciones de un componente o el error esperado de las estimaciones, con una probabilidad de ocurrencia dada.

Es posible detectar con lo anterior, las estimaciones particulares que resultan significativamente diferentes al grueso de estimaciones y, por ende, determinar las estimaciones atípicas dentro de una colección, asumiendo como tales a aquéllas que presenten desviaciones, por encima del margen de error tolerado. Esto supone, considerar como típica una estimación por el hecho de que cumpla una condición dada y, como atípica si no la cumple.

Lo anterior no puede presuponer nada más allá del propio carácter atípico, asignado a la estimación, pues asumir que ello es indicativo de un determinado error esperable, respecto a cualquier parámetro externo a las propias mediciones, como es el resultado registrado, implica considerar como dada una distribución normal de las diversas estimaciones colectadas y la inexistencia de un sesgo implícito que afecte al conjunto de mediciones, aspectos que tendrían que validarse. Además, el hecho mismo de detectar una estimación como atípica, no puede derivar en su exclusión de un cálculo de tendencia central para un conjunto de estimaciones dadas con el fin de supuestamente robustecerlo, sin evaluar antes la pertinencia del procedimiento de exclusión y, en su caso, constatar el cumplimiento de otras condiciones demandas conforme los diversos criterios de eliminación de casos extremos disponibles en la literatura. Lo que es más: es deseable analizar la relación empírica entre el cumplimiento de la condición de haber sido una estimación típica o atípica, respecto a un conjunto de estimaciones y la de estar dentro o fuera del margen de error, respecto al resultado para interpretar el significado de la condición de atipicidad de una estimación.

Regresando a la argumentación central, y partiendo de la definición del error esperado para una estimación, es posible calcular un error normalizado de dicha estimación, respecto al dato real, como el cociente del error absoluto observado entre el error esperado, que correspondería al estimador de la magnitud relativa del sesgo con sentido más comúnmente empleado en estadística. La generación de este estimador, demanda la disposición de un dato adicional para cada unidad analizada, a los requeridos para el cálculo del error en estimadores previamente disponibles: el tamaño de la muestra o, en su caso, la proporción de casos asignados a algún contendiente respecto del total de casos observados. Empero, este aumento de la información demandada para el cálculo, se premia con la disposición de un estimador que no se afecta sustancialmente por el número de contendientes involucrados en cada elección.

De hecho, al amortiguar los efectos relacionados con el peso relativo de las proporciones correspondientes, al darles una ponderación ajustada a una relación cuadrática, permite que la eventual agregación o diferenciación de contendientes menores, no propicie una importante sobreestimación o subestimación del error calculado; esto es: se logra disponer de un estimador del

error que no sea tan sensible a los valores extremos. Por ende, su cálculo no solo puede realizarse para todo contendiente del que exista un reporte diferenciado de proporciones observadas, sino que su estimación resulta coherente y comparable sin importar el peso relativo del componente. Es posible calcular el error normalizado de la diferencia entre las estimaciones para un par de componentes respecto a los valores realmente dados. Esto resulta del encuentro de la desviación estándar de la estimación de la diferencia, que supone considerar para el cálculo del error, en una diferencia los errores en las estimaciones para cada una de las proporciones consideradas.

Cabe recordar que la desviación de una muestra particular puede afectarse por el diseño adoptado. Técnicas de estratificación o división de la población de estudio en grupos que se suponen homogéneos con respecto a alguna característica a estudiar, suelen aumentar la precisión, misma que disminuye cuando se recurre a la selección de conglomerados, como es usual y por lo general necesario, por razones prácticas en encuestas electorales. Cabe mencionar que los métodos directos para computar las varianzas de muestras complejas son tediosos, por lo que suelen usarse procedimientos abreviados que no miden la incertidumbre sobre sesgos o errores de respuesta y, por lo tanto, son estimaciones del error mínimo involucrado en la medición.

Para medir el efecto del diseño se calcula el número de casos bajo un diseño de muestra aleatoria simple que sería equivalente a la muestra actual bajo el diseño muestral específico, adoptado. En el caso de diseños de muestras complejas, el tamaño real de la muestra se determina multiplicando el tamaño efectivo de la muestra por el efecto de diseño, entendido como la razón de la varianza verdadera de un estadístico construido, tomando en cuenta el diseño muestral respecto de la varianza para una muestra aleatoria simple, con el mismo número de casos.

Simplificando procedimientos, lo relevante en el caso de encuestas electorales es el tamaño de la muestra asignada que corresponde al número de casos disponibles para lograr una precisión con un muestreo aleatorio simple. Con base en esto último número de casos ajustado por el efecto de diseño, es posible establecer un error esperado para la muestra asignada equivalente a un muestreo aleatorio simple o bien, un margen de error esperado para la muestra asignada incluyendo el efecto de diseño para hacerla equivalente a un muestreo

aleatorio simple, con base en lo cual puede establecerse el intervalo de la estimación para la muestra asignada, incluyendo el efecto de diseño, equivalente a un muestreo aleatorio simple.

Luego, se puede calcular el error normalizado de dicha estimación respecto al dato real como el cociente del error absoluto estimado entre el error esperado, dada la muestra efectivamente asignada. Éste es el estadístico más exacto para el cálculo del error normalizado de una estimación que permite establecer su magnitud y sentido. Basados en él, debiera analizarse la existencia, o no, de un sesgo adicional al error aleatorio esperado en la medición. Sin embargo, cabe repetir, su cálculo lleva incorporado el efecto del diseño como parte del error estocástico, por lo que la adopción de este estimador para fines comparativos entre estudios, tendería a premiar los diseños menos eficientes y castigar aquéllos que hubieran tenido mayor eficiencia en su arranque. Además, es de precisar que los efectos de diseño son diferentes para distintos subgrupos poblacionales y respuestas a reactivos, por lo que no existe un efecto genéricamente aplicable a la totalidad de datos, producto de una encuesta.

Es por ello que resulta preferible, en aras de la comparabilidad entre estudios, y adicionado al problema de disposición de los datos requeridos para una estimación más exacta para cada estudio, establecer un estimador del error normalizado de las mediciones que dé cuenta del error estadístico, obtenido cual si se tratara de muestreos aleatorios simples en todos los casos, como procedimiento para dotar de uniformidad a la base para la estimación y, dado que es el método con restricciones mínimas, usarlo como punto de referencia para evaluar la eficiencia de los diseños y la presencia potencial de sesgos adicionales. Retomando entonces el estimador orientado de la magnitud relativa del sesgo normalizada, se enfrenta nuevamente el problema de que su suma aritmética no permite la estimación de un valor que dé cuenta de la magnitud agregada del error, pues solamente arrojaría un dato relacionado con el tamaño de las salientes en un sentido u otro que fuera predominante en un conjunto de estimaciones dado. Por ello, puede adoptarse la estrategia de calcular la distancia euclidiana de los errores normalizados, idealmente calculando la raíz cuadrada de la mitad de la suma de los cuadrados de los errores normalizados, para evitar la duplicidad de contabilización de errores, dado que las salientes que se tengan provocan entrantes de similar magnitud.

Sin embargo y en aras de lograr la mayor simplicidad, la nueva propuesta de estimador del error entre las proporciones arrojadas por una encuesta y las proporciones reales (E), parte precisamente de obtener los valores absolutos de los errores normalizados de la estimación para cada contendiente y luego extraer su promedio. Haciendo esto, y debido al comportamiento parabólico del error esperado para un tamaño de muestra dado, se disminuye de manera significativa, el efecto que puede tener el número o tamaño de los componentes considerados para fines de estimación del error medio, siendo poco sensible a decisiones de separación o agrupamiento de contendientes por problemas en los reportes de resultados. Este estimador adquiere valores que van de cero, cuando existe perfecta coincidencia entre las proporciones observadas y las reales, hasta cerca de infinito e incluso, adquiere los mismo valores que el cálculo de la distancia euclidiana cuando se calcula para los datos de elecciones con dos contendientes, aunque difiere cuando los contendientes son más de dos. Asimismo, como siempre, es viable calcular el error absoluto normalizado de la estimación de la diferencia entre dos contendientes cualesquiera, lo que en el caso de los dos componentes mayores dará cuenta del error normalizado para la estimación del margen de victoria.

¿Qué ocurre si aplicamos este estimador de la (in)exactitud de las encuestas al conjunto de mediciones seleccionadas que forman la colección de estudios previos a las elecciones presidenciales de México durante el presente siglo? Lo primero que se ve es que, si bien existen cambios respecto al cálculo convencional a través de M3, estos no resultan mayores. En 2000 el promedio de encuestas estuvo a 1.7 desviaciones estándar en promedio y la media de errores estándar por encuesta fue de 2.4; para 2006, en promedio las encuestas tuvieron un error estandarizado de 1.8 y cada medición tuvo, promediando, 2.0 desviaciones estándar respecto del resultado; finalmente, para 2012, la encuesta de encuestas como forma de agregación habría estado desviada 1.9 errores estándar y cada medición habría estado 2.2 desviaciones estándar en promedio distante del resultado.

Aunque en apariencia, los datos anteriores no difieren tanto de los anteriormente vistos, sí hay un hecho que destacar: 2012 no resulta peor que 2000, sino que incluso puede afirmarse que las mediciones finales fueron un poco más

certeras en términos de desviaciones estándar y vistas como conjunto que lo observado doce años antes.

CUADRO 7

ERROR ESTANDARIZADO DE LAS ENCUESTAS FINALES EN 2000

Encuestas finales 2000	Muestra efectiva	ERROR CONTABILIZADO				E
		VFQ	FLO	CCS	Otros	
Encuestadora						
ARCOP	1204	-0.7	0.9	-0.3	0.2	0.51
GEA-ISA	1481	-2.4	1.0	1.0	1.5	1.47
Democracy Watch	1496	-2.0	-0.7	2.9	0.9	1.62
GAUSSC	1337	-2.2	1.2	2.0	-2.2	1.88
Demotecnia	1150	0.3	-2.1	-0.9	4.9	2.05
Alduncin y Asociados	1251	-3.3	3.7	-1.0	0.8	2.18
Grupo Reforma	1316	-3.8	4.4	-0.9	0.0	2.27
AC Nielsen	2016	-4.1	4.6	-1.2	1.1	2.76
MERCAEI	1472	-4.3	3.2	2.0	-1.6	2.79
Consultores y MP	838	-1.6	3.9	-2.1	-3.7	2.83
CEO	1676	-2.5	-2.0	3.4	3.9	2.94
Zogby	1399	-3.8	4.8	-2.3	1.7	3.12
Pearson	2205	-4.4	5.5	-2.2	1.4	3.38
MUND	1253	-6.2	-0.5	7.7	-1.2	3.92
Promedio de encuestas	1546	-3.0	2.1	0.8	0.9	1.69
Resultado oficial		2.9	2.6	2.2	1.9	2.41

Fuente: elaboración propia con base en datos de *El papel de las encuestas en las elecciones federales. Memorias del Taller Sumiya 2000*, 2001.

CUADRO 8

ERROR ESTANDARIZADO DE LAS ENCUESTAS FINALES EN 2006

Encuestas finales 2006	Muestra efectiva	ERROR CONTABILIZADO				M3
		FCH	RMP	AMLO	Otros	
Encuestadora		FCH	RMP	AMLO	Otros	M3
GEA-ISA	1403	0.8	0.0	-0.2	-1.8	0.71
ARCOP	1204	0.1	1.6	-1.7	0.4	0.93
Grupo Reforma	1013	-0.1	2.0	-1.6	-0.5	1.06
Consultores y MP	1721	-2.3	2.2	0.1	0.5	1.26
El Universal	1220	-2.1	2.4	-0.2	0.4	1.28
CEO	823	-2.4	2.6	0.4	-1.3	1.68
Consulta Mitofsky	1032	-1.8	2.4	-1.4	1.9	1.85
BGC Ulises Beltrán	1590	-3.3	3.6	-0.2	0.4	1.89
Parametría	1738	-3.0	2.2	-0.4	3.0	2.15
GAUSSC	870	-1.2	3.3	-3.4	2.7	2.64
Alduncin y Asociados	1733	-1.7	5.5	-3.8	0.4	2.86
Zogby	990	-3.1	3.6	-1.9	2.9	2.88
Demotecnia	1240	-4.9	5.1	-0.7	1.2	2.96
INDEMERC	2990	-2.6	7.0	-3.0	-2.9	3.90
Promedio de encuestas	1398	-2.0	3.2	-1.3	0.8	1.82
Resultado oficial		2.1	3.1	1.4	1.4	2.00

Fuente: cálculos del autor con base en la información contenida en: *Memorias del seminario: encuestas y elecciones 2006*, Instituto Federal Electoral-Asociación Mexicana de Agencias de Investigación de Mercado y Opinión Pública-Asociación Mundial de Investigadores de la Opinión Pública-Consejo de Investigadores de la Opinión Pública, México, 2010.

CUADRO 9

ERROR ESTANDARIZADO DE LAS ENCUESTAS FINALES EN 2012

107

Encuestas finales 2012	Muestra efectiva	ERROR CONTABILIZADO				M3
		JVM	EPN	AMLO	GQT	
Encuestadora						
MERCAEI	899	0.7	-0.4	-0.9	1.7	0.94
Covarrubias y Asociados	1410	-0.1	0.6	-2.0	3.3	1.48
Demotecnia	648	-0.9	2.5	-1.6	-1.0	1.49
Grupo Reforma	975	-2.4	0.6	0.0	3.3	1.58
Berumen y Asociados	864	-1.4	3.1	-1.9	-0.6	1.77
IPSOS-BIMSA	1293	-1.8	1.3	-1.1	3.1	1.82
Consulta Mitofsky	996	-0.8	3.1	-3.1	1.3	2.06
BGC Ulises Beltrán	984	3.3	-0.1	-3.8	1.3	2.14
ARCOP	810	-1.7	2.7	-2.3	2.2	2.23
Parametría	964	-2.8	4.8	-2.7	-0.2	2.61
Buendía & Laredo	2791	-4.7	2.7	1.8	-1.5	2.67
Con Estadística	943	-1.0	3.2	-4.0	2.9	2.77
GEA-ISA	1752	-1.7	4.9	-4.2	1.0	2.94
INDEMERC	1805	-3.3	6.8	-5.1	1.5	4.18
Promedio de encuestas	1224	-1.2	2.6	-2.3	1.6	1.93
Resultado oficial		1.9	2.6	2.5	1.8	2.19

Fuente: elaboración propia con base en datos del Séptimo Informe que presenta la Secretaría Ejecutiva al Consejo General del Instituto Federal Electoral, respecto del cumplimiento del Acuerdo CG411/2011, Instituto Federal Electoral, México, 26 de julio de 2012.

¿Qué pasa si nuevamente intentamos una agregación de las (in)exactitudes observadas en una única tabla que dé cuenta de los resultados obtenidos a lo largo del siglo por las diversas firmas encuestadoras, pero esta vez a través del estadístico generado como alternativa y no del convencional? Lo primero que observamos es que el vicio en el cálculo propiciado por la diferente cantidad de mediciones presentadas por cada firma encuestadora, tiende a disminuir, ubicándose entre los primeros lugares del ordenamiento casas que han reportado varias veces datos previos a las elecciones. En segundo lugar, que la mejor estimación con este método corresponde a la firma que queda en el primer sitio,

manteniéndose la peor observación al final del cuadro. Además, dado que lo estadísticamente esperable es que el error medio medido como desviaciones estándar se ubique dentro de un límite preestablecido (dos desviaciones entre estimación y resultado), es posible separar a las firmas encuestadoras cuyas mediciones se han ubicado dentro de lo esperable de aquellas que se ha desviado más.

CUADRO 10

INEXACTITUD PROMEDIO OBSERVADA ENTRE LAS ESTIMACIONES POR ENCUESTA Y LOS RESULTADOS OFICIALES DE LAS ELECCIONES PRESIDENCIALES DE MÉXICO CONFORME AL MEDIDOR E (2000-2012)

ORDEN	ENCUESTADORA	CASOS	ELECCIÓN PRESIDENCIAL			E
			2000	2006	2012	
1	ARCOP	3	0.51	0.93	2.14	1.19
2	El Universal	1		1.28		1.28
3	Covarrubias y Asociados	1			1.48	1.48
4	IPSOS-BIMSA	1			1.49	1.49
5	GEA-ISA	3	1.47	0.71	2.61	1.60
6	MERCAEI	2	2.27		0.94	1.60
7	Democracy Watch	1	1.62			1.62
8	Grupo Reforma	3	2.18	1.26	1.82	1.76
9	Consulta Mitofsky	2		1.89	1.77	1.83
10	Consultores y MP	2	2.79	1.06		1.92
11	Parametría	2		1.68	2.23	1.96
12	BGC Ulises Beltrán	2		1.85	2.06	1.96
13	Demotecnia	3	2.05	2.96	1.58	2.20
14	Berumen y Asociados	1			2.67	2.67
15	Zogby	2	2.83	2.64		2.74
16	CEO (U de G)	2	3.38	2.15		2.76
17	AC Nielsen	1	2.76			2.76
18	Con Estadística	1			2.77	2.77

19	GAUSSC	2	1.88	3.90		2.89
20	Alduncin y Asociados	2	2.94	2.86		2.90
21	Buendía & Laredo	1			2.94	2.94
22	Pearson	1	3.12			3.12
23	INDEMERC	2		2.88	4.18	3.53
24	MUND	1	3.92			3.92
PROMEDIO			2.41	2.00	2.19	2.20

Fuente: elaboración propia con base en los datos compilados en los cuadros 3a, 3b y 3c

En promedio, el cambio de orden de las diversas casas encuestadoras es de apenas dos y medio lugares, aunque una gran parte de los cambios se debe a variaciones en la colocación de seis de las 24 firmas consideradas. Este hecho de confianza en una afirmación de cierre a este apartado: pareciera que sí existen mecanismos sólidos que den cuenta de la exactitud de las encuestas electorales y que la historia que puede contarse tiene algún sentido y sustento lógico y técnico.

De los cuadros 2d y 3d puede también extraerse otra conclusión: no existe una correlación clara ni definida entre la exactitud lograda por una firma encuestadora en una ocasión y la conseguida en la subsecuente ocasión, cuando se ha presentado. Firmas con datos muy próximos al resultado en una estimación dada, previa a una elección, se han alejado en forma significativa en la oportunidad posterior, y viceversa. Así, los antecedentes de exactitud o desviación no garantizan el “éxito” esperado en un momento posterior. Solamente la agregación de experiencia de las casas encuestadoras, incluso más allá de los eventos de medición de procesos de elección presidencial, puede ir otorgando el *expertis* requerido para intentar evitar sesgos e incrementar la probabilidad, siempre incierta, de tener mediciones que se aproximen a la realidad.

CONCLUSIÓN

Al final de cuentas, sí existe una mayor o menor cercanía cuantificable de las encuestas preelectorales con los resultados oficiales. El empleo de este procedimiento no solamente responde a esta lógica, sino a razones empíricas, dada la recurrente coincidencia entre lo medido y el resultado oficial, que alimenta la expectativa depositada en el instrumento como supuesto anticipador del

reparto de votos. Es primordialmente por eso que los medios las difunden y que el público las atiende, aunque ello no sea, del todo, científicamente sustentable. La importancia otorgada a estos contrastes tiene que ver, además, con la relevancia pública de los estudios, su indemostrado impacto en el electorado y la supuesta disposición de un parámetro contra el cual contrastar. Negarlo mediante el recurso de que el único estimador pertinente del verdadero parámetro poblacional, sería el promedio de los estimadores centrales de las repetidas mediciones que se realicen y que la (in)exactitud de una estimación particular debiera calcularse, por la diferencia, contra la media del conjunto de estimaciones disponibles, no solo es una falacia, sino una forma torpe de eludir el problema central que es la demanda pública de que las estimaciones preelectorales, producto de encuestas aproximen al elector al conocimiento de lo factible en la jornada electoral.

Y ello nos lleva a un último punto que no debiéramos dejar de lado: si bien los errores estimados parecieran encontrarse dentro de un rango que ha derivado en una “normalidad”, en promedio, menos de la mitad de las estimaciones por componente (candidato), se han encontrado dentro del margen esperable de dos desviaciones estándar. Esto es: la mitad de las veces, por ponerlo de manera sencilla, el dato atribuido a un contendiente se encuentra fuera de lo estadísticamente esperado, si el origen del error fuera muestral (basado en el método aleatorio simple). Esto es válido para el conjunto de mediciones durante el presente siglo y para cada una de las elecciones presidenciales que se han celebrado.

Se podrá alegar, en favor de las encuestadoras que es imposible recurrir, hoy día en México, a procedimientos aleatorios simples, por lo que siempre se debiera añadir un inevitable efecto de diseño. Se podrá argumentar que existen muy diversas fuentes de error no muestral, como el fraseo de preguntas, efectos no anticipados de preguntas previas o posteriores a otras preguntas en el cuestionario, y contingencias que se presentan durante el levantamiento de campo que pueden generar sesgos, incluso de carácter no sistemático, que afecten la calidad de la estimación. Esto es conocido y reconocido por los profesionales del campo.

CUADRO 11

COMPARACIÓN ENTRE LAS FRECUENCIAS OBSERVADAS Y TEÓRICAS

111

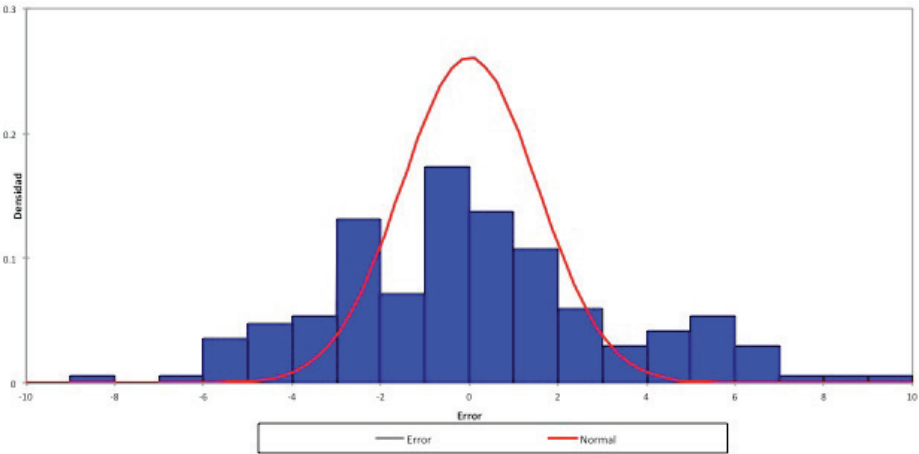
Límite inferior (DE)	Límite superior (DE)	Frecuencia (componentes)	Frecuencia (distribución normal)	%
-10	-9	0	0	
-9	-8	1	0	
-8	-7	0	0	
-7	-6	1	0	28%
-6	-5	6	0	
-5	-4	8	1	
-4	-3	9	3	
-3	-2	22	12	
-2	-1	12	27	
-1	0	29	41	
0	1	23	41	49%
1	2	18	27	
2	3	10	12	
3	4	5	3	
4	5	7	1	
5	6	9	0	
6	7	5	0	23%
7	8	1	0	
8	9	1	0	
9	10	1	0	

Fuente: Cálculos del autor con base en los datos del Cuadro 10.

GRÁFICA 1

HISTOGRAMA DE DISTRIBUCIÓN DE DESVIACIONES ESTÁNDAR, OBSERVADAS ENTRE ESTIMACIONES POR ENCUESTA Y RESULTADOS OFICIALES EN MÉXICO (2000-2012) Y CURVA DE DISTRIBUCIÓN NORMAL

112



Fuente: elaboración propia.

Pero el hecho está allí: la distribución conforme a las desviaciones estándar de los datos reportados en las estimaciones por encuesta en fecha próxima a los comicios y los resultados oficiales de los mismos, distan de distribuirse de manera normal. No solo se alejan en demasía, en promedio, sino que tienden a conformar una curva con colas excesivamente pronunciadas. Así, lo correcto, lo profesional, sería que cada casa encuestadora enfrentara la tarea no solamente de advertir los riesgos de sesgos que pueden producirse por razones no estadísticas, sino de reconocer el efecto de diseño que pueda detectar en sus datos y determinar el sesgo que haya observado en sus mediciones públicas anteriores y que, eventualmente, pudieran afectar la calidad de sus estimaciones actuales.

BIBLIOGRAFÍA

113

Crespi, Irving (1988). *Pre-Election Polling: Sources of Accuracy and Error*, Russell Sage Foundation, Nueva York.

Instituto Federal Electoral (2001), *El papel de las encuestas en las elecciones federales. Memorias del Taller Sumiya 2000*, Instituto Federal Electoral-Asociación Mexicana de Agencias de Investigación de Mercado y Opinión Pública-Colegio Nacional de Actuarios, México.

_____ (2010), *Memorias del seminario: encuestas y elecciones 2006*, Instituto Federal Electoral-Asociación Mexicana de Agencias de Investigación de Mercado y Opinión Pública-Asociación Mundial de Investigadores de la Opinión Pública-Consejo de Investigadores de la Opinión Pública, México.

_____ (2012), *Séptimo informe que presenta la Secretaría Ejecutiva al Consejo General del Instituto Federal Electoral respecto del cumplimiento del Acuerdo CG411/2011, por el que se establecen los lineamientos así como los criterios generales de carácter científico que deberán observar las personas físicas y morales que pretendan ordenar, realizar y/o publicar encuestas por muestreo, encuestas de salida y/o conteos rápidos durante el Proceso Electoral 2011-2012*, México, 26 de julio.

Instituto Nacional Electoral (2014), *Acuerdo INE/CG220/2014 del Consejo General del Instituto Federal Electoral por el que se establecen los lineamientos así como los criterios generales de carácter científico que deberán observar las personas físicas y morales que pretendan ordenar, realizar y/o publicar encuestas por muestreo, encuestas de salida y/o conteos rápidos que tengan por fin dar a conocer preferencias electorales, así como preferencias sobre consultas populares, durante los procesos electorales federales y locales*, México, 22 de octubre.

Martin, Elizabeth A.; Michael W. Traugott y Courtney Kennedy (2005). "A Review and Proposal for a New Measure of Poll Accuracy", en *Public Opinion Quarterly*, vol. 69, núm. 3, pp. 342-369.

Mitofsky, Warren (1998). "Review: Was 1996 a Worse Year for Polls Than 1948?", en *The Public Opinion Quarterly*, vol. 62, núm. 2, pp. 230-249.

Mosteller, Frederick (1949). "Measuring the error", en: Mosteller, Frederick; Herbert Hyman; Philip J. McCarthy, *et. al.*, *The Pre-election Polls of 1948, Report of the Committee on Analysis of Pre-election polls and forecasts*, Bulletin 60, Social Science Research Council, Nueva York.

114

Recibido: 24/05/2015

Aceptado: 15/06/2015