

# Las elecciones de 2006

## *un análisis del conteo rápido*

**P**ara estimar los resultados electorales se han realizado numerosos y muy diversos ejercicios estadísticos, desde encuestas preelectorales, hasta encuestas de salida, conteos rápidos y encuestas postelectorales. En general, están basados en muestras probabilísticas y producen estimaciones en distintos tiempos y con diferentes grados de incertidumbre, que son inherentes a cualquier valor estimado partiendo de una muestra probabilística. También conllevan la incertidumbre propia de la calidad del proceso de recolección y procesamiento de la información.

Las encuestas preelectorales toman la información de votantes potenciales, generalmente ciudadanos que dicen tener credencial para votar. Además de estimar los porcentajes de votación en favor de los distintos partidos, candidatos o fuerzas contendientes, obtienen información adicional que permite a los diversos partidos construir un perfil del votante potencial, así como aquella de interés para quien planea las campañas preelectorales y de publicidad. Los resultados difundidos tienen cierto tiempo de desfase respecto al tiempo en que se recolecta, pues se pone a disposición del público una vez que la información fue levantada, filtrada y procesada. La metodología es variada, no sólo en cuanto al diseño muestral usado, que depende de la forma en que se construye, y del tamaño y la distribución de la muestra, sino también en cuanto a los mé-

todos de recolección en el campo y los métodos de estimación.

Un tipo particular de encuestas preelectorales son las telefónicas, las cuales proporcionan resultados que reflejan la opinión de sólo una parte de la población, si consideramos que la penetración del teléfono en el país es de aproximadamente 37%, y va de acuerdo con las estimaciones de la muestra asociada al Censo General de Población y Vivienda 2000 de 12% en el estado de Chiapas a 67% en el Distrito Federal.

Las encuestas de salida toman la información de votantes consumados; esto es, se selecciona una muestra, en principio probabilística, de personas que han votado, las cuales son entrevistadas al salir de la casilla. La muestra probabilística la conforman las casillas donde se realizan las entrevistas. Estas encuestas producen estimaciones durante el día de la elección en diversos cortes hasta terminar al momento del cierre de casillas, a las 18 horas del tiempo local.

Los conteos rápidos se realizan la misma noche de la elección y toman la información del resultado de la casilla reportado, sea en la hoja de operaciones de cómputo o en el acta de escrutinio y cómputo aprobada por los funcionarios de casilla. Al igual que en el caso anterior, las casillas seleccionadas donde se recoge la información constituyen



Guillermina Eslava

una muestra probabilística. El producto de un conteo rápido generalmente estima el resultado final de los cómputos distritales del 5 de julio, hasta antes de considerar imputaciones y ajustes finales realizados y aprobados por el Tribunal Electoral del Poder Judicial de la Federación (TEPJF).

Finalmente, las encuestas postelectorales son un ejercicio que se realiza después de la elección y que recoge la información de ciudadanos para dar estimaciones y perfiles de los votantes consumados, y de los no votantes, de acuerdo con su opinión o inclinación hacia los distintos partidos.

Todos estos ejercicios, basados en muestras probabilísticas, se realizan en diversos niveles: municipal, distrital, estatal y nacional. Algunos son difundidos en los medios de comunicación y otros son empleados para monitoreo interno de los partidos o grupos de interés.

Existen otros dos ejercicios que no están basados en muestras, sino que se enfocan al registro de los resultados por casilla para el total de las casillas, son el Programa de resultados electorales preliminares (PREP) y los cómputos distritales.

El PREP registra, de forma electrónica y secuencial, los resultados de cada casilla electoral tomando el dato que

aparece en la hoja de escrutinio y cómputo destinada específicamente al programa, y entregada por el funcionario de casilla en la oficina de cómputo del distrito que le corresponde. Inicia el registro de la información el día de la elección al cierre de las casillas y ofrece resultados desde las 20 horas del tiempo del centro del país.

Sin embargo, el resultado oficial se basa en los cómputos distritales de todas las actas de casilla realizados el 5 de julio en cada uno de los 300 distritos electorales. Finalmente, las cifras reportadas por estos cómputos sufren cambios o ajustes derivados de las impugnaciones, correcciones o cancelaciones aprobadas por el TEPJF. Así, puede decirse que el resultado de los cómputos distritales de actas, iniciados el 5 de julio y reportados el día 6, también aproximan el resultado final oficial emitido por el TEPJF. La fuente de información en este cómputo oficial es el acta de escrutinio y cómputo de cada casilla electoral.

Los ejercicios de conteo rápido para estimar, la noche de la elección, los porcentajes de votos en favor de cada partido contendiente, y por lo tanto del candidato ganador, son realizados con cierta regularidad, en los niveles nacional, estatal, distrital y municipal, por instituciones gubernamentales, no gubernamentales, empresas privadas y medios de comunicación.

Para las elecciones presidenciales de 2006 se efectuaron varios conteos rápidos, algunos públicos y otros no, con una diversidad de estrategias en cuanto al diseño muestral y al tamaño de muestra. Este tipo de ejercicios inició en 1994 para las elecciones presidenciales con la intención de ofrecer resultados casi inmediatos a la sociedad. En ese año, fueron comisionadas por las autoridades federales tres empresas independientes para realizar tres conteos rápidos, y los resultados no se presentaron al público. En las elecciones federales de 2000, nuevamente fueron comisionadas tres empresas independientes; pero esa vez se les encomendó recolectar la información en las casillas electorales correspondientes a tres muestras independientes y un comité técnico se encargó de procesar la información y obtener los valores estimados de los porcentajes de votación en favor de cada partido o coalición; en aquella ocasión, los resultados fueron difundidos por el IFE la misma noche de la elección. En 2006, por primera vez el IFE realiza un conteo rápido de forma independiente y autónoma; todo el proceso lo realizó personal interno bajo la coordinación de un Comité técnico asesor para los conteos rápidos (CTACR), el cual se encargó de diseñar la muestra y de procesar la información recabada la noche de la elección.



A pesar de que se obtuvieron estimaciones puntuales de los porcentajes, dos de los intervalos de confiabilidad asociados no se separaron por más de 0.006, margen acordado para poder proporcionar los resultados la misma noche. Por esta razón no se difundieron las cifras estimadas, sino solamente el aviso de que estadísticamente no era posible distinguir con un alto grado de confiabilidad a un candidato ganador. Las estimaciones se hicieron públicas el día 3 de julio en conferencia de prensa.

Desde el punto de vista metodológico, el conteo rápido es un ejercicio que, basado en una muestra probabilística de resultados de casillas electorales, estima el porcentaje de votos en favor de cada uno de los partidos o coaliciones contendientes. Esta estimación lleva asociada un error acotado en términos probabilísticos.

$$P(|\hat{P}_A - P_A| < \delta) = 1 - \alpha$$

donde  $\hat{P}_A$  es la estimación de  $P_A$ ;  $P_A$  es la proporción de votos en favor de un partido o coalición A;  $\delta$  es la cota del error y  $\alpha$  es un valor entre 0 y 1 que determina el nivel de confianza o el grado de incertidumbre de la estimación. El error depende directamente de la variabilidad de la característica de interés que pretende medirse en las unidades poblacionales, e inversamente del tamaño de la muestra.

Para determinar el diseño y la estimación de la cota al error, se hizo uso de los resultados por casilla registrados en el PREP de las elecciones presidenciales del año 2000 —estos valores estimados aparecen en el cuadro 4. En la figura 1 se presentan gráficamente la información del PREP 2000 y el desarrollo de las estimaciones clásicas del conteo rápido del mismo año.

Por su parte, las principales características del conteo rápido de 2006 son las siguientes: un marco muestral de 130 500 casillas contenidas en 300 distritos electorales y listadas en el padrón electoral; el diseño muestral es un muestreo estratificado simple de casillas; el tamaño y la distribución de muestra fue de 7 636 casillas en 481 estratos; el criterio de estratificación fue distrito y tipo de sección —urbana, no urbana—; la información en el campo fue recopilada por capacitadores asistentes electorales del IFE; la unidad de observación fue la hoja de operaciones de cómputo de la casilla; la transmisión telefónica se efectuó hacia un centro de acopio distrital y luego, en forma electrónica, hacia una central de cómputo, donde se realizó el procesamiento, el análisis y la producción de las estimaciones.

## ELECCIONES PRIMARIAS



Fue un proceso largo, complejo y un trabajo de equipo coordinado. La información inicial viaja y se procesa para llegar hasta la conformación de una estimación final; en forma esquemática puede verse como sigue: voto → urna → conteo de votos → hoja de operaciones de cómputo → oficina distrital vía telefónica → oficina central vía electrónica → computadoras personales → procesamiento de la información y generación de estimadores → análisis de consistencia y cotejo de resultados para alcanzar una conclusión → comunicación de resultados.

El conteo rápido, por ser de carácter institucional y cuyos resultados serían la base de un comunicado difundido a la ciudadanía la misma noche de la elección, estuvo sujeto a acuerdos previos y a restricciones como guardar la confidencialidad de la muestra; realizar estimaciones por intervalos de cuando menos una confiabilidad de 95%; que para identificar un partido o coalición ganadora, los intervalos asociados a los porcentajes estimados de votos para los partidos mayoritarios deberán distar en al menos 0.6% (0.006); el ejercicio se realizaría entre las 18 y las 22:30 o 24 horas; y respetar el formato de comunicación de resultados aprobado por los consejeros electorales y los representantes de los partidos.

### Los resultados numéricos

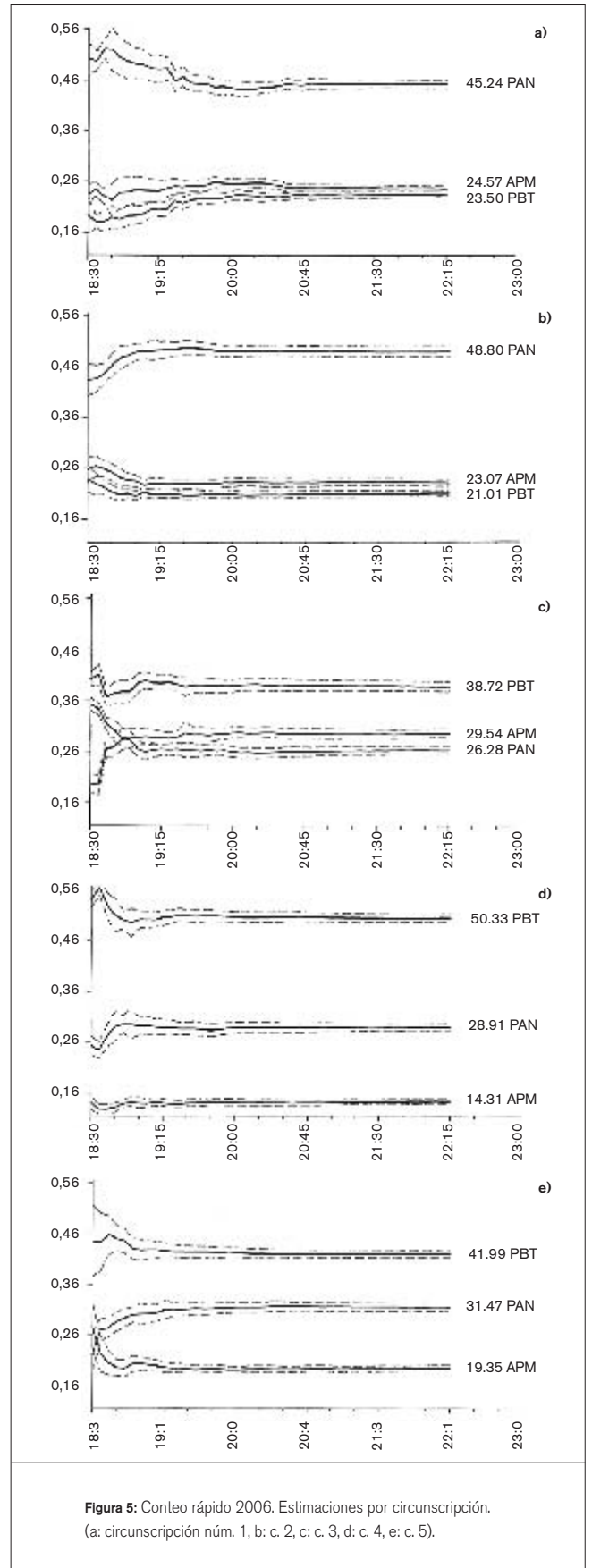
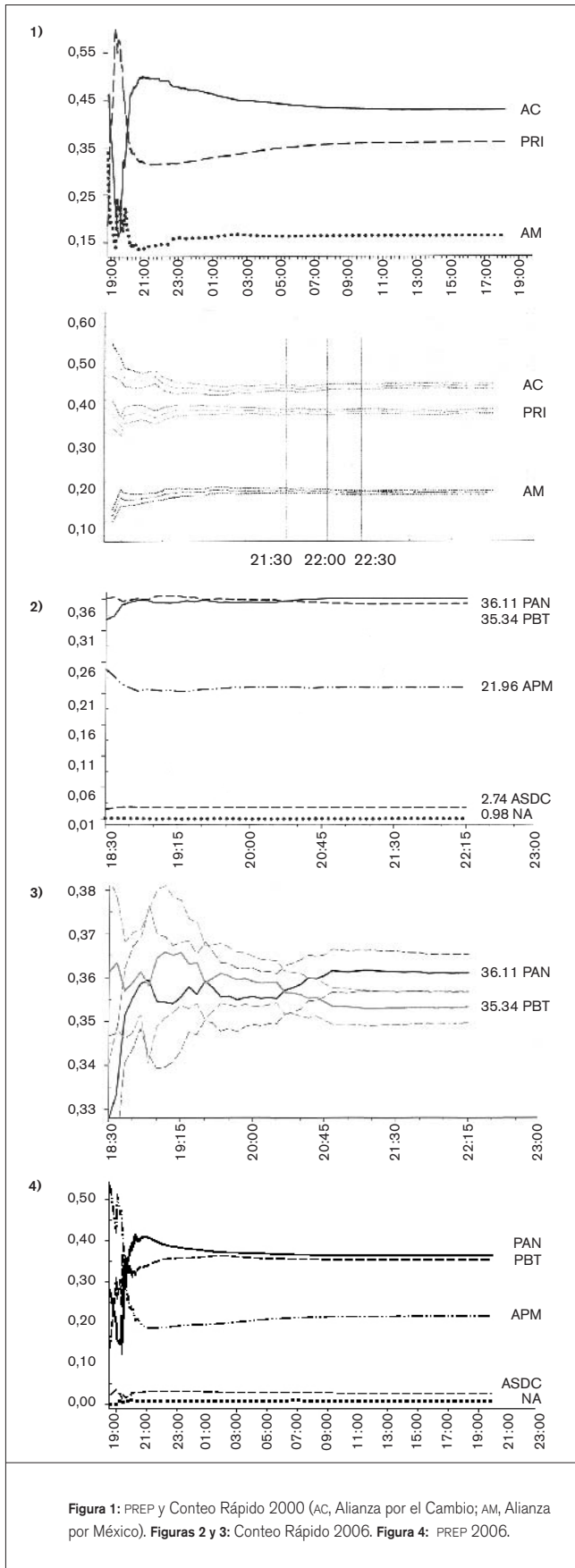
Las estimaciones presentadas se obtuvieron con tres métodos: el clásico, el bayesiano y el robusto. El cálculo se realizó con base en la información recibida en la remesa de

1)		PAN	APM	CPBT	NA	ASDC
	TERPJF	35.89	22.23	35.33	0.96	2.71
	CÓMPUTOS DISTRITALES	35.89	22.26	35.31	0.96	2.70
	ESTIMADOR CLÁSICO PUNTUAL	36.11	21.96	35.34	0.98	2.74
	CLÁSICO	(35.68, 36.53)	(21.66, 22.26)	(34.97, 35.70)	(0.93, 1.03)	(2.6, 2.8)
	BAYESIANO	(35.77, 36.40)	(21.72, 22.24)	(35.07, 35.63)	(0.94, 1.05)	(2.6, 2.8)
	ROBUSTO	(35.25, 37.40)	(20.85, 22.70)	(34.24, 36.38)	(0.75, 1.19)	(2.4, 3.13)

2)	CIRCUNSCRIPCIONES		PAN		APM		CPBT		NA		ASDC	
			%	INTERVALO	%	INTERVALO	%	INTERVALO	%	INTERVALO	%	INTERVALO
	Circ. 1	C.R.	45.24	(44.42, 46.06)	24.57	(23.96, 25.17)	23.50	(22.92, 24.09)	1.25	(1.14, 1.35)	2.89	(2.76, 3.02)
		C.DIST.	45.03		24.79		23.55		1.20		2.85	
	Circ. 2	C.R.	48.80	(47.85, 49.75)	23.07	(22.38, 23.76)	21.01	(20.42, 21.59)	1.13	(1.02, 1.23)	2.67	(2.53, 2.8)
		C.DIST.	48.29		23.56		21.05		1.15		2.64	
	Circ. 3	C.R.	26.28	(25.47, 27.1)	29.54	(28.71, 30.37)	38.72	(37.79, 39.66)	0.69	(0.5, 0.87)	1.58	(1.48, 1.67)
		C.DIST.	25.82		29.89		38.87		0.61		1.52	
	Circ. 4	C.R.	28.91	(28.11, 29.7)	14.31	(13.84, 14.78)	50.33	(49.53, 51.13)	0.81	(0.74, 0.89)	3.14	(3.04, 3.24)
		C.DIST.	29.05		14.64		49.08		0.82		3.12	
	Circ. 5	C.R.	31.47	(30.67, 32.28)	19.35	(18.81, 19.88)	41.99	(41.25, 42.74)	1.04	(0.97, 1.11)	3.36	(3.23, 3.49)
		C.DIST.	31.51		19.39		42.05		1.03		3.30	

3)		PAN		APM		CPBT		NA		ASDC		
		%	INTERVALO	%	INTERVALO	%	INTERVALO	%	INTERVALO	%	INTERVALO	
	28 DISTRITOS CON ALTA PRE- SENCIA INDÍGENA	C.R.	37.24	(36.84, 37.64)	20.91	(20.64, 21.19)	35.23	(34.9, 35.57)	0.97	(0.93, 1.01)	2.88	(2.83, 2.94)
		C.DIST.	37.02		21.17		35.23		0.97		2.84	
	272 DISTRITOS CON BAJA O SIN PRESENCIA INDÍGENA	C.R.	23.23	(21.81, 24.64)	33.86	(32.3, 35.43)	36.52	(34.73, 38.32)	1.11	(0.67, 1.56)	1.16	(1.04, 1.27)
		C.DIST.	23.28		34.39		36.16		0.90		1.10	

4)	PARTIDO O COALICIÓN 2000/2006	ERROR ESPERADO APROXIMADO $\delta_{PREP\ 2000} = 3 \cdot \sqrt{V(\hat{P}_A)}$ n = 7636	ERROR OBSERVADO ESTIMADO $\delta_{C.R.\ 2006} = 3 \cdot \sqrt{\hat{V}(\hat{P}_A)}$ n = 7263	ERROR ESPERADO (A POSTERIORI) $\delta_{C.DIST.\ 2006} = 3 \cdot \sqrt{V(\hat{P}_A)}$ n = 7636	
	ACAM/ PAN		0.39	0.42	0.38
	PRI/ APM		0.34	0.30	0.28
	AMEX/ CPBT		0.28	0.36	0.34
	Errores multiplicados por 100.				
	<b>Cuadro 1</b>	Elecciones 2006. Resultados y estimaciones.			
	<b>Cuadro 2</b>	Elecciones 2006. A nivel circunscripciones.			
	<b>Cuadro 3</b>	Elecciones 2006. A nivel distritos con alta presencia indígena.			
	<b>Cuadro 4</b>	Elecciones 2006. Errores de estimaciones poblacionales y muestrales.			



las 22:15 horas, donde se tenía resultados de 7 263 casillas distribuidas en 480 estratos y en los 32 estados —estos datos se presentan en el cuadro 1, que además muestra los resultados procesados a partir de la base de datos de los cómputos distritales, que esta disponible en forma electrónica en el portal del IFE.

En las figuras 2 y 3 se ofrecen las gráficas que ilustran la evolución de las estimaciones para llegar hasta las estimaciones presentados en el cuadro 1. En la figura 4 se presenta la evolución del registro secuencial de resultados en el PREP. Queremos comentar que los valores observados en las gráficas 2 y 3 correspondientes a horas tempranas, no pueden considerarse como estimaciones específicas, básicamente por insuficiencia de la muestra. Sin embargo, el historial de las estimaciones para llegar al resultado reportado sirve de guía para tomar decisiones sobre la suficiencia de la muestra y la estabilidad en las estimaciones.

#### La riqueza del conteo rápido 2006

El diseño muestral del conteo rápido, incluyendo el tamaño de la muestra, permitió generar estimaciones estadísticamente confiables para subpoblaciones de interés. Por ejemplo, para la subpoblación contenida en los 28 distri-

tos con alta presencia indígena, y para cada una de las cinco circunscripciones. La primera circunscripción corresponde a Baja California, Baja California Sur, Chihuahua, Durango, Jalisco, Nayarit, Sinaloa y Sonora; la segunda, Aguascalientes, Coahuila, Guanajuato, Nuevo León, Querétaro, San Luis Potosí, Tamaulipas y Zacatecas; la tercera, Campeche, Chiapas, Oaxaca, Quintana Roo, Tabasco, Veracruz y Yucatán; la cuarta, el Distrito Federal, Guerrero, Morelos, Puebla y Tlaxcala; y la quinta, Colima, Hidalgo, Estado México y Michoacán. El diseño muestral usado fue un diseño autoponderado, con ajustes ligeros para evitar estratos muy pequeños, lo cual hace que la información sea robusta cuando se emplean distintas técnicas de estimación, además, el tamaño y la distribución de la muestra fue lo suficientemente grande para poder obtener estimaciones en niveles más desagregados de la población. En los cuadros 2 y 3 se presenta información desagregada con estimaciones generadas bajo el método clásico.

En el cuadro 2 puede observarse las estimaciones generadas en el nivel de circunscripción, junto con los valores de los cómputos distritales del 5 de julio —esto se muestra gráficamente en la figura 5, mientras que en el cuadro 3 se presentan las estimaciones puntuales y por intervalo para la subpoblación conformada por los 28 distritos electorales con alta presencia de población indígena, y para su complemento contenido en los 272 distritos restantes.

Finalmente, aunque quizá sólo de interés metodológico, en el cuadro 4 presentamos los errores esperados aproximados, éstos son valores estimados en base con los resultados por casilla reportados en el PREP de 2000, y son aproximados porque existió una redistribución entre 2000 y 2006, por lo que el marco muestral o padrón electoral de 2000 difiere del correspondiente a 2006. Estos errores los contrastamos con los errores estimados observados el día de la elección, basados en la muestra del conteo rápido de 2006. Por último, una vez pasada la elección, y haciendo uso de los resultados por casillas de los cómputos distritales, se evalúa el error esperado, considerando el diseño muestral usado en el conteo rápido de 2006.

#### Comentarios

En general, la realización del conteo rápido resultó exitoso y generó estimaciones muy cercanas a las reportadas con base en los registros del conteo distrital de actas, como puede constatarse al comparar los valores estimados con los valores reportados.



Es de esperarse que además del error de estimación del conteo rápido, que se acota desde el diseño bajo una medida de probabilidad y no de certeza, existen discrepancias respecto al resultado final oficial emitido por el TEFPJ, pues en esta instancia ocurren modificaciones al resultado emitido por los conteos distritales, las cuales son producto, como ya se señaló, de las modificaciones derivadas de impugnaciones, anulaciones o correcciones en los resultados de las actas.

Por otro lado, es un ejercicio estadístico que enriquece las formas de registro y de estimación para elecciones futuras, dejando constancia de un procedimiento con resultados obtenidos tanto de forma global como desagregada y que, junto con los registros electrónicos del PREP y de los conteos distritales, son de gran utilidad, no sólo para la academia, sino también para la ciudadanía interesada en ello, y desde luego para las instituciones encargadas de conducir las elecciones en cualquier nivel en el país. ❄



#### Guillermina Eslava

Facultad de Ciencias,  
Universidad Nacional Autónoma de México.

#### AGRADECIMIENTOS

Todo este trabajo es parte, como ya lo indiqué, del proceso preparado y generado por el CTACR, su coordinador y personal del IFE, a quienes expreso mi gratitud por la colaboración de equipo para la obtención de resul-

tados. El Comité técnico asesor para los conteos rápidos estuvo conformado como sigue: Act. M. Cervera Flores (IF), Dra. G. Eslava Gómez (FC), Dr. R. Hernández Cid (ITAM), Dr. I. Méndez Ramírez (IMAS), Dr. M. Mendoza Ramírez (ITAM); y fue coordinado por el Dr. A. Alonso y Coria (IFE). Agradezco también al Dr. R. Mena Chávez (IMAS), quien elaboró el programa de cómputo usado la noche de la elección para la generación de una de las tres estimaciones clásicas. Las gráficas presentadas en este trabajo fueron elaboradas por Gonzalo Pérez de la Cruz, pasante de actuaría de la Facultad de Ciencias, a quien agradezco su atenta colaboración. Final-

mente agradezco al Dr. Ramón Peralta y Fabi, director de esta Facultad, por el apoyo durante el semestre que participe en el CTACR y la invitación para escribir este trabajo.

#### IMÁGENES

P. 31: "La próxima elección", en *El Ahuizote Jacobino*, núm. 20, 1904. P. 32: "Pasaje bíblico nacional"; p. 33: "Elecciones primarias"; p. 36: "Rapacidad castigada", en *El Hijo del Ahuizote*, 1903 y 1902. P. 37: "Orgías reeleccionistas", en *El Colmillo Público*, núm. 147, 1906.

**Palabras clave:** Elecciones presidenciales 2006, Conteo rápido, Diseño muestral

**Key words:** Presidential elections 2006, Quick count, Sample design

**Resumen:** En este espacio presentamos de manera breve algunos elementos del conteo rápido del 2006, pero sobre todo, los valores estimados obtenidos la noche de la elección y una serie de figuras que ilustran gráficamente algunos aspectos de los resultados numéricos obtenidos.

**Abstract:** We briefly introduce some of the elements of the 2006 *quick* counting, but above all we show the estimated results obtained the night of the elections and a series of figures that show graphically some of the features of the numerical results obtained.

Guillermina Eslava Gómez es Actuaría y Maestra en estadísticas por la UNAM, y Doctora en estadística por la Universidad de Oxford, Inglaterra. Actualmente es Profesora de Tiempo Completo de la Facultad de Ciencias, UNAM.

**Recepción:** 11 de septiembre de 2006, aceptado 15 de septiembre de 2006