

UNA OPINION SOBRE EL PREP Y EL CONTEO DISTRITAL

Víctor Romero Rochín

Investigador Titular del Instituto de Física.

UNAM

Después de analizar las bases de datos del Programa de Resultados Electorales Preliminares (PREP) y del Conteo Distrital (CD), se puede afirmar que existe una posibilidad estadísticamente insoslayable de que se haya interferido con el sistema de cómputo del IFE que realizó la contabilidad de los votos reportados en las casillas electorales. Inmediatamente aclaramos que esto no implica que el IFE esté involucrado: bien puede suceder que el IFE sea tan víctima de la posible interferencia, como el resto de nosotros.

La razón fundamental para tal afirmación es que la evolución del resultado electoral, conforme se iban contabilizando las casillas, tiene un comportamiento que estadísticamente es prácticamente imposible. Es muy importante señalar que el estudio no pretende afirmar que el *resultado* final es incorrecto. De ninguna manera. La intención de este estudio es sólo hacer señalamientos sobre la *forma* tan poco típica en la que evolucionó el proceso de la contabilización de los votos.

Veamos los detalles, primero el PREP y luego el CD.

PREP

El proceso de captura de los datos asentados en las actas de las casillas el día de la elección, inició oficialmente a las 6 PM en 300 distritos electorales repartidos en el país.

En cada distrito se tienen 10 capturistas, lo que arroja 3000 individuos capturando las actas de las casillas electorales de manera *simultánea*. Esto nos indicaría que estamos frente a un proceso con una muy fuerte componente azarosa. Eso, sin considerar la inevitable presencia de errores humanos.

Al mismo tiempo, nos afirman los estudios geográficos y sociológicos, existen factores que nos indican que las capturas en los diferentes distritos no ocurren de manera homogénea. Tomando esto cuenta, se esperaría que al principio de la captura de la votación, digamos el primer 5 % de las casillas, y al final, el último 5%, sí hubiera una diferencia en la com-

posición estadística de esas muestras con respecto al restante 90%. Lo robusto de la muestra del 90% del total, en el sentido estadístico, lo podemos afirmar pues el número de votos es enorme, 42 millones.

Combinando los dos puntos anteriores se esperaría que del 10% al 90%, la composición porcentual en cada muestra de, digamos, 1000 casillas, fuera la misma. Es decir, que en una gran parte del proceso, el azar predominara sobre los cambios sistemáticos. Esto lleva a la conclusión de que el porcentaje del voto de cada candidato debería estabilizarse alrededor del resultado final, no más tarde del 70% de los votos. Es esta la razón por la cual podemos creerle a un conteo rápido o a una encuesta. Siempre y cuando, por supuesto, la diferencia porcentual entre los candidatos sea mayor al error del conteo o de la encuesta.

Los resultados del PREP, no sólo se alejan de lo esperado arriba mencionado sino que, más alarmante, muestran tendencias *ordenadas* en el tiempo. Decimos que su comportamiento es estadísticamente atípico, que sólo quiere decir que su probabilidad de ocurrencia es muy baja.

Veamos las Figuras 1 y 2. La Fig. 1 muestra el porciento del voto *acumulado*, con respecto al número de casillas contabilizadas, de los partidos PAN y PRD, y la Fig. 2 los correspondientes al PRI. Es importante entender qué es un resultado “acumulado”: quiere decir que cuando se tenían 40,000 casillas contadas, el cálculo del porcentaje se hace con esas 40,000; y cuando ya se tienen 50,000, el porcentaje se hace ahora con las 50,000. Este porcentaje debe tender a estabilizarse pues cada vez se tienen mas votos. Es de notarse que en el PREP, el porcentaje de los tres contendientes principales, PAN, PRI y PRD, *nunca* se estabilizó. Es decir, si observamos a cada candidato por separado, en ningún momento podemos decir cuál podría ser el resultado final.

Para ver con más detalle lo peculiar del resultado anterior, lo que tenemos que analizar son los votos que le llegan a cada candidato, digamos *cada* 1000 casillas, y que corresponde a aproximadamente a 500 mil votos. Lo que queremos analizar es, entonces, *muestra por muestra* en el orden en el que llegaron, y no el acumulado. Lo que se espera, por la fuerte componente azarosa explicada arriba, es que las muestras sean estadísticamente independientes. Es decir, una no puede tener ninguna conexión de *orden* con respecto a la siguiente o a la que esté mucho más

adelante o mucho más atrás.

La Figura 3 muestra el porcentaje de votos por *cada* 1000 casillas conforme fueron llegando. El comportamiento de 0 a 20,000 puede ser entendible por razones geográficas y sociológicas. De 20,000 a 60,000 el PAN y el PRD parecen tener comportamiento típico en el que se observa que suben y bajan sus votos pero alrededor de un valor fijo. Y de 60,000 a 117,000 existe un comportamiento muy atípico ya que, no sólo aparece un cambio abrupto en el comportamiento, sino que lo hacen de manera *ordenada*; vea la Figura 4 donde aparecen los datos sólo del PAN y del PRI, y donde se muestra el número de votos absolutos y no el porcentaje: note que uno baja y el otro sube *en orden* aunque, recuerde, que cualquier par de puntos consecutivos corresponden a dos muestras de 1000 casillas independientes una de la otra. Es muy peculiar que el PRI siempre sube de votos absolutos, no porcentaje, en muestras de 1000 casillas consecutivas, o sea, que aparecen ordenadas de menor a mayor. Al final llegan las últimas 10,000 casillas en las que los votos se reparten equitativamente entre los tres candidatos: otro hecho sorprendente en sí.

De la Figura 3 note otro aspecto al que hay que ponerle atención: *Ningún* conjunto de puntos de los 3 partidos principales, y que corresponden a muestras de aproximadamente 500 mil votantes, refleja el resultado final del PREP que fue de 36.38 % PAN, 35.34% PRD y 21.57% PRI. Es decir, las encuestas previas a la votación, así como las de salida, se acercaron más al resultado final, con menos de 5,000 personas, que 117 muestras de 500,000 votantes.

Las conclusiones de los puntos anteriores son (1) la falta de estabilización del porcentaje en el voto acumulado, (2) la diferencia tan grande, entre las muestras de 1000 casillas independientes cada una, y el resultado final; y (3) el *orden* en el que llegaron los votos a cada candidato *por cada* 1000 casillas. Los puntos (1) y (2), aunque estadísticamente son muy atípicos en sí, podrían justificarse apelando a que eventos improbables sí ocurren. Sin embargo, la conclusión (3) definitivamente no. Esta última requiere de una sincronización casi perfecta para ordenar la llegada de los resultados ... siendo que, estaremos de acuerdo todos, nadie puede predecir el futuro para arreglar el orden en el que deben suceder las cosas.

Y la última explicación que no podemos descartar es que, debido a que toda la información se centralizó en las computadoras del IFE, un agente externo interfirió con la presentación de los resultados. A su vez, esta conclusión nos lleva a desconfiar de la veracidad del proceso de captura de los datos y no del proceso electoral en sí.

CONTEO DISTRITAL

El CD se llevó a cabo el día miércoles 5 de julio con una mecánica diferente al del día de las elecciones, pero muy importante, se contabilizó con el mismo sistema de cómputo y de informática del IFE. Adelantamos que las observaciones hechas al PREP se repiten aquí.

Debido, pues, a que nuevamente la información se capturó en 300 distritos que supuestamente no estuvieron comunicándose entre sí, podemos suponer que se tiene una fuerte componente de azar. Sin embargo, se ha reportado que en casillas donde se tenía una preferencia clara por el PAN, la revisión minuciosa de las actas resultó en un retraso del envío de los resultados de dichas casillas; este es un hecho que debe tomarse en cuenta.

La Figura 5 muestra el porcentaje del voto *acumulado* del PAN y del PRD, conforme se fueron contabilizando las casillas. Tres aspectos son peculiares de esta figura.

La primera es que, como en el PREP, nunca se estabilizó el porcentaje de ninguno de estos dos candidatos, aunque sí fue así con los demás partidos; es curioso ver que en el último 0.01% de contabilidad de los votos hubo todavía un cambio a favor del PAN.

El segundo punto es que las curvas se cruzan; en general, si el proceso fuera muy azaroso, las curvas no se cruzarían, a menos que los porcentajes fueran muy cercanos, pero en ese caso se esperarían varios cruces. La explicación que se ha dado es el hecho arriba mencionado, es decir, que los votos de casillas favorables al PAN llegaron después de los votos de las casillas con mayoría del PRD. Eso explicaría el cruce de las curvas pero, como ahondaremos con más detalle más adelante, no explica la *forma* en la que lo hizo.

La tercera observación es que las curvas aparecen a la vista como si una fuera esencialmente el espejo de la otra. Esto puede explicarse verificando que los votos que “perdía” uno lo “ganaba” el otro, sin “darle”

nada a ninguno de los otros candidatos en la contienda; y parece ser que este es el caso ya que los otros candidatos muestran una variación mínima en sus porcentajes. Es decir, los otros ahora sí se estabilizaron pero nunca lo hicieron el PRD y el PAN, como si la contienda fuera entre ellos nada más. Una regularidad para resaltar.

Para tener una mejor comprensión del “espejo” y del cruce tan tardío, analizamos ahora el arribo de la captura de los votos cada 1000 casillas aproximadamente, como lo hicimos con el PREP. La Figura 7 muestra el resultado del análisis. Primero vemos que el PRI se mantiene, de manera estadísticamente típica, alrededor de su valor final y la “competencia” se da sólo entre el PAN y el PRD. Claramente corroboramos que, efectivamente, las casillas con mayoría del PAN son capturadas después que aquellas con mayoría del PRD, y eso explica el cruce. Lo que es nuevamente estadísticamente muy improbable es que el arribo de las casillas con preferencia para el PAN lo hacen de manera *ordenada* en el tiempo. Vea el detalle del 80% en adelante de la Figura 7. En ese intervalo, prácticamente en todas las casillas gana el PAN, pero lo que queremos resaltar, una vez más, es que primero llegan las casillas en las que la diferencia es nula o muy pequeña, luego las que tienen una diferencia un poco más grande, luego mayor, y así sucesivamente hasta que al final llegan las casillas en las que el PAN aventaja de manera máxima ... pero ¿cómo puede ocurrir si no hay relación entre una muestra y la siguiente, ni comunicación simultánea en aproximadamente 150 distritos electorales? Bajo el supuesto que el conteo estuvo bien hecho, uno esperaría que hubieran llegado “revueltas” y que, por lo tanto, el cruce en el porcentaje acumulado hubiera ocurrido antes y que los porcentajes se hubieran estabilizado.

Las conclusiones son esencialmente las mismas que en el caso del PREP. E insistimos una vez más que lo que provoca mayor sorpresa es el orden, ascendente o descendente, del número de votos conforme se contabilizan nuevas casillas, siendo que las muestras son independientes unas de las otras. Note que el resultado final no cambiaría si hubieran llegado al azar o en cualquier otra secuencia. Un orden de esta naturaleza, de nuevo, no puede descartarse en términos estadísticos, aunque si así fuera tendría una probabilidad increíblemente pequeña. La se-

gunda opción otra vez, es una sincronización rayando en lo perfecto de 150 distritos electorales repartidos en todo el país y un análisis sobre la marcha de cuál acta tiene la mayor diferencia que las otras, etc. Y la posibilidad final, que no puede ni debe descartarse a la ligera, es la intervención de un agente externo al sistema de cómputo del IFE.

Es irónico que lo que se podrían llamar “irregularidades” en este caso, son las regularidades observadas en un proceso que se espera tenga una fuerte componente azarosa.

Como mencionamos desde el inicio de este reporte, las observaciones aquí vertidas nos llevan inevitablemente a considerar como una verdadera posibilidad una injerencia en el sistema de cómputo del IFE y, por lo tanto, que para salir de dudas es sano realizar un conteo adicional de los votos y/o las actas de cada casilla. Y como una condición necesaria, evitando que la información se vuelva a centralizar en las mismas computadoras.