

## **COMITÉ TÉCNICO ASESOR DEL CONTEO RÁPIDO COTECORA**

Informe Final sobre el  
Conteo Rápido de la Elección para  
Gobernador del Estado de México 2017

**26 de junio de 2017**

## Control de versiones

Versión	Descripción	Fecha
1.0	Elaboración	23 JUNIO 2017
1.1	Revisión	26 JUNIO 2017

## ÍNDICE DE CONTENIDO

I. INTRODUCCIÓN.....	4
II. ANTECEDENTES .....	4
III. MARCO JURÍDICO .....	6
IV. OBJETIVO DEL COTECORA.....	13
V. OBJETIVO DE LOS CONTEOS RÁPIDOS.....	13
VI. PROGRAMA DE ACTIVIDADES .....	14
VII. SESIONES DEL COTECORA .....	16
VIII. CRITERIOS CIENTÍFICOS.....	17
IX. SIMULACROS DEL CONTEO RÁPIDO EN EL ESTADO DE MÉXICO .....	40
X. SELECCIÓN DE LA MUESTRA. ....	44
XI. DIA DE LA JORNADA ELECTORAL. ....	45
XII. ANÁLISIS POST-JORNADA ELECTORAL.....	53
XIII. CONCLUSIONES Y RECOMENDACIONES.....	54

## I. INTRODUCCIÓN

La definición de los conteos rápidos la podemos encontrar en el primer inciso artículo 356 del Reglamento de Elecciones aprobado por el Consejo General del INE INE/CG/661/2016.

*Los conteos rápidos son el procedimiento estadístico diseñado con la finalidad de estimar con oportunidad las tendencias de los resultados finales de una elección, a partir de una muestra probabilística de resultado de actas de escrutinio y cómputo de las casillas electorales, cuyo tamaño y composición se establecen previamente, de acuerdo a un esquema de selección específico de una elección determinada y cuyas conclusiones se presentan la noche de la jornada electoral.*

Si bien el resultado de la elección se determina considerando el cómputo definitivo de los votos y resolviendo las impugnaciones o controversias presentadas en la elección, el conteo rápido ofrece una estimación de los resultados con alta precisión estadística y esa estimación se comunica a la población durante la noche del mismo día de la elección. Estas dos características fortalecen la confianza de todos los ciudadanos sobre los procedimientos y la validez de los resultados, asimismo, permite inhibir pronunciamientos de diferentes actores no necesariamente basados en la realidad de los resultados.

Para llevar a cabo este ejercicio, la propia normativa antes citada establece en su numeral 362, la conveniencia y necesidad de contar con Comité Técnico Asesor de los Conteos Rápidos (COTECORA), para brindar al IEEM asesoría para el diseño, implementación y operación de los conteos. En ese tenor, el 14 de diciembre de 2016 fue aprobado por el Consejo General del IEEM el acuerdo IIEEM/CG/115/2016 *Por el que se integra el Comité Técnico Asesor de los Conteos Rápidos para el Proceso Electoral 2016-2017.*

Ahora bien, para llevar su desarrollo a buen fin, es necesario contar con una ruta clara, reflejada en instrumentos, transparentes, que generen confianza entre los distintos actores de la contienda electoral.

## II. ANTECEDENTES

- 1.- El siete de septiembre de dos mil dieciséis, el Consejo General del Instituto Electoral del Estado de México, celebró sesión solemne por la que se dio inicio formalmente al Proceso Electoral Ordinario 2016-2017, para elegir al Gobernador Constitucional del Estado de México, para el periodo comprendido del 16 de septiembre de 2017 al 15 de septiembre de 2023.
- 2.- En sesión extraordinaria del siete de septiembre de dos mil dieciséis, el Consejo General del Instituto Nacional Electoral aprobó el Acuerdo

INE/CG661/2016, denominado: “Acuerdo del Consejo General del Instituto Nacional Electoral por el que se aprueba el Reglamento de Elecciones del Instituto Nacional Electoral”, mismo que fue publicado en el Diario Oficial de la Federación, el trece de septiembre del mismo año.

- 3.- El doce de septiembre de dos mil dieciséis, la H. “LIX” Legislatura Local, expidió el Decreto número 124, publicado en el Periódico Oficial del Gobierno del Estado Libre y Soberano de México, “Gaceta del Gobierno”, en la misma fecha, en donde se convoca a los ciudadanos del Estado de México y a los partidos políticos con derecho a participar en la elección ordinaria para elegir al Gobernador Constitucional del Estado de México, para el periodo comprendido del 16 de septiembre de 2017 al 15 de septiembre de 2023.
- 4.- En sesión ordinaria del veintidós de septiembre del dos mil dieciséis, el Órgano Superior de Dirección de este Instituto, aprobó el Acuerdo IEEM/CG/83/2016, por el que creó la Comisión Especial para la atención del Programa de Resultados Electorales Preliminares y Conteos Rápidos; la cual quedó integrada de la siguiente forma: Presidenta, Consejera Electoral Mtra. Palmira Tapia Palacios; como integrantes, los Consejeros Electorales, Mtro. Miguel Ángel García Hernández y Mtro. Saúl Mandujano Rubio; como Secretario Técnico el Secretario Ejecutivo y con un representante de cada partido político con registro ante este Instituto.

En el Considerando XXVII, se estableció entre otros aspectos, como motivo de creación y dos de sus objetivos, los siguientes:

#### **“Motivo de Creación**

*Contar con una Comisión que auxilie al Consejo General en la coordinación de las actividades relativas a la implementación y operación del Programa de Resultados Electorales Preliminares (PREP), así como para la realización de los conteos rápidos, en la elección de Titular del Poder Ejecutivo del Estado de México.*

#### **Objetivos**

- *Proponer al Consejo General la designación de los órganos internos responsables de la coordinación de las actividades del Programa de Resultados Electorales Preliminares (PREP) y de los Conteos Rápidos.*
- ...
- *Supervisar la correcta integración y funcionamiento del Comité Técnico Asesor del Programa de Resultados Electorales Preliminares (COTAPREP) y Conteos Rápidos (COTECORA).*
- ...”

- 5.- En sesión extraordinaria celebrada el ocho de diciembre de dos mil dieciséis, la Comisión Especial para la atención del Programa de Resultados Electorales Preliminares y Conteos Rápidos, aprobó el Acuerdo IEEM/CEPREP y CR/03/2016 por el que propone a este Consejo General, la integración del Comité Técnico Asesor de los Conteos Rápidos (COTECORA); y;
- 6.- El 14 de diciembre de 2016 fue aprobado por el Consejo General del IEEM el acuerdo IEM/CG/115/2016 Por el que se integra el Comité Técnico Asesor de los Conteos Rápidos para el Proceso Electoral 2016-2017.
- 7.- El día 16 de mayo de 2017 fueron aprobados por el Consejo General del IEEM los acuerdos IEM/CG/120/2017, IEM/CG/121/2017 los Criterios Científicos, Protocolo para la selección y resguardo de la muestra para la realización del Conteo Rápido y la Guía de Procedimiento de la operación logística de los conteos rápidos en el ámbito distrital para la elección de Gobernador en el Proceso Electoral 2016-2017.

### **III. MARCO JURÍDICO**

- I. El artículo 41, párrafo segundo, Base V, de la Constitución Política de los Estados Unidos Mexicanos, establece que la organización de las elecciones es una función estatal que se realiza a través del Instituto Nacional Electoral y de los Organismos Públicos Locales.

Por su parte, la base antes citada, en su Apartado B, inciso a), numeral 5, menciona que para los procesos electorales federales y locales, corresponde al Instituto Nacional Electoral en los términos que establecen la propia Constitución y las leyes, entre otros aspectos, lo relativo a las reglas, lineamientos, criterios y formatos en materia de conteos rápidos.

Asimismo, el Apartado C, párrafo primero, numeral 8, de la base referida, determina que en las entidades federativas, las elecciones locales estarán a cargo de Organismos Públicos Locales en los términos que señala la propia Constitución, que ejercerán entre otras funciones, las relacionadas a los resultados preliminares y conteos rápidos, conforme a los lineamientos establecidos en el Apartado B, de la propia base.

- II. El artículo 32, numeral 1, inciso a), fracción V, de la Ley General de Instituciones y Procedimientos Electorales, en adelante Ley, establece como atribución del Instituto Nacional Electoral para los procesos electorales federales y locales, entre otras, el establecimiento de las reglas, lineamientos, criterios y formatos en materia de conteos rápidos.

- III.** El artículo 98, numerales 1 y 2, de la Ley, señala que los Organismos Públicos Locales están dotados de personalidad jurídica y patrimonio propios y gozarán de autonomía en su funcionamiento e independencia en sus decisiones, en los términos previstos en la Constitución General, la propia Ley, las constituciones y leyes locales; asimismo, que son autoridad en la materia electoral, en los términos que establece la Constitución Política de los Estados Unidos Mexicanos, la referida Ley y las leyes locales correspondientes.
- IV.** De conformidad con el artículo 104, numeral 1, incisos a), f) y n) de la Ley, corresponde a los Organismos Públicos Locales:
- Aplicar las disposiciones generales, reglas, lineamientos, criterios y formatos que, en ejercicio de las facultades le confiere la Constitución Federal y dicha Ley, establezca el Instituto Nacional Electoral.
  - Llevar a cabo las actividades necesarias para la preparación de la jornada electoral.
  - Ordenar la realización de conteos rápidos basados en las actas de escrutinio y cómputo de casilla a fin de conocer las tendencias de los resultados el día de la jornada electoral, de conformidad con los lineamientos emitidos por el Instituto Nacional Electoral.
- V.** El artículo 220, numerales 1 y 2, de la Ley, establece que el Instituto Nacional Electoral y los Organismos Públicos Locales determinarán la viabilidad en la realización de los conteos rápidos; y que de igual manera, las personas físicas o morales que realicen estos conteos pondrán a su consideración, las metodologías y financiamiento para su elaboración y términos para dar a conocer los resultados de conformidad con los criterios que para cada caso se determinen.
- VI.** El artículo 1º, numerales 1 y 2, del Reglamento de Elecciones del Instituto Nacional Electoral, en adelante Reglamento, establece que el mismo tiene por objeto regular las disposiciones aplicables en materia de instituciones y procedimientos electorales, así como la operación de los actos y actividades vinculados al desarrollo de los procesos electorales que corresponde realizar, en el ámbito de sus respectivas competencias, al Instituto Nacional Electoral y a los Organismos Públicos Locales de las entidades federativas; además de que su observancia es general y obligatoria para el Instituto Nacional Electoral, los Organismos Públicos Locales de las entidades federativas, en lo que corresponda; los partidos políticos, precandidatos, aspirantes a candidatos independientes, candidatos, así como para las personas físicas y morales vinculadas a alguna etapa o procedimiento regulado en dicho ordenamiento.

- VII.** El artículo 355, numeral 1, del Reglamento, refiere que las disposiciones contenidas en el Capítulo III, del Título III, del mismo, son aplicables para el Instituto Nacional Electoral y los Organismos Públicos Locales, en sus respectivos ámbitos de competencia, respecto de todos los procesos electorales federales y locales que celebren, y tienen por objeto establecer las directrices y procedimientos a los que deben sujetarse dichas autoridades para el diseño, implementación, operación y difusión de la metodología y resultados de los conteos rápidos.
- VIII.** El artículo 356, numerales 1 al 4, del Reglamento, señala que:
- Los conteos rápidos son el procedimiento estadístico diseñado con la finalidad de estimar con oportunidad las tendencias de los resultados finales de una elección, a partir de una muestra probabilística de resultados de actas de escrutinio y cómputo de las casillas electorales, cuyo tamaño y composición se establecen previamente, de acuerdo a un esquema de selección específica de una elección determinada, y cuyas conclusiones se presentan la noche de la jornada electoral.
  - En el diseño, implementación y operación de los conteos rápidos, las autoridades electorales y el Comité Técnico de la materia, deberán garantizar la seguridad, transparencia, confiabilidad, certeza, calidad e integridad del procedimiento estadístico, así como el profesionalismo y la máxima publicidad en la ejecución de sus trabajos.
  - El procedimiento establecido por las autoridades electorales y el Comité Técnico de la materia, garantizará la precisión, así como la confiabilidad de los resultados del conteo rápido, considerando los factores que fundamentalmente se relacionan, por una parte, con la información que emplean y, por otra, con los métodos estadísticos con que se procesa esa información.
  - El objetivo del conteo rápido es producir estimaciones por intervalos del porcentaje de votación para estimar la tendencia en la elección, el cual incluirá además la estimación del porcentaje de participación ciudadana.
- IX.** Que como lo dispone el artículo 357, numeral 1, del Reglamento, el Consejo General del Instituto Nacional Electoral y los Órganos Superior de Dirección de los Organismos Públicos Locales, tendrán la facultad de determinar la realización de los conteos rápidos en sus respectivos ámbitos de competencia; así como, que cada Organismo Público Local, en su caso, informará al Consejo General del Instituto Nacional Electoral sobre su determinación dentro de los tres días posteriores a que ello ocurra, por conducto de la Unidad Técnica de Vinculación con Organismos Públicos Locales.

Asimismo, el numeral 2, del mismo artículo, menciona que no obstante, los Organismos Públicos Locales deberán realizar conteos rápidos en el caso de elecciones de Gobernador o de Jefe de Gobierno en el caso de la Ciudad de México.

- X.** Que el artículo 358, numeral 1, del Reglamento, prevé que la Presidencia del Consejo General del Instituto Nacional Electoral o del Órgano Superior de Dirección de los Organismos Públicos Locales, en su respectivo ámbito de competencia, será el responsable de coordinar el desarrollo de las actividades de los conteos rápidos.

Por su parte, el numeral 2, de dicho artículo, señala que tanto el Instituto Nacional Electoral como los Organismos Públicos Locales, en su ámbito de competencia, son responsables de la asignación de los recursos humanos, financieros y materiales para la implementación de los conteos rápidos.

- XI.** Que el artículo 359, del Reglamento, refiere que los Órganos Superiores de los Organismos Públicos Locales, resolverán los aspectos no previstos en el Capítulo III, del Título III del ordenamiento en cita, apegándose a las disposiciones legales y los principios rectores de la función electoral.
- XII.** Que como lo señala el artículo 361, numerales 1 y 2, del Reglamento, el Instituto Nacional Electoral y los Organismos Públicos Locales podrán contratar a personas físicas o morales para que los apoyen en las actividades de los conteos rápidos, que consideren necesarias, las cuales deberán respetar las directrices establecidas en el Capítulo III del Título III del propio Reglamento, así como los acuerdos de contratación que aprueben los Órganos Superior de Dirección respectivos. Asimismo, que las personas físicas o morales contratadas para apoyar en actividades operativas de los conteos rápidos, no podrán participar en el diseño y selección de la muestra, ni en la difusión de la metodología y los resultados.
- XIII.** Que en términos del artículo 362, numeral 1, del Reglamento, el Consejo General del Instituto Nacional Electoral y su homólogo en los Organismos Públicos Locales, dentro de su ámbito de competencia, deberán aprobar, al menos cuatro meses antes de la fecha en que deba celebrarse la respectiva jornada electoral, la integración de un Comité Técnico Asesor que les brindará asesoría para el diseño, implementación y operación de los conteos rápidos, cuyos integrantes deberán iniciar sus funciones al día siguiente de su designación.

Asimismo, el numeral 2, del mismo artículo señala que el Comité Técnico Asesor de los Conteos Rápidos se integrará por Asesores Técnicos: de tres a cinco expertos en métodos estadísticos y diseño muestral, con derecho a voz y voto; y un Secretario Técnico: funcionario del Instituto Nacional Electoral o del Organismo Público Local según corresponda, con derecho a

voz, quien será el enlace entre el comité y el Consejo General respectivo y auxiliará en todo momento a los asesores técnicos.

**XIV.** Que el artículo 364, del Reglamento, dispone que el Comité Técnico Asesor de los Conteos Rápidos deberá celebrar una sesión de instalación, de acuerdo a lo siguiente:

- a) El Secretario Técnico convocará a los miembros del Comité Técnico Asesor de los Conteos Rápidos e invitará a los integrantes del Consejo General u Órgano Superior de Dirección del Organismo Público Local, según corresponda, así como a los representantes de los partidos políticos y, en su caso, de candidaturas independientes, con una anticipación de por lo menos setenta y dos horas.
- b) En el orden del día de dicha sesión, se presentará para su aprobación, el proyecto del plan de trabajo y el calendario de sesiones ordinarias. Asimismo, se darán a conocer los objetivos, las dinámicas de trabajo, así como la normatividad para la realización de los conteos rápidos.

**XV.** Que como lo prevé el artículo 365, numeral 1, del Reglamento, las sesiones del Comité Técnico Asesor de Conteos Rápidos podrán ser ordinarias y extraordinarias:

- a) Serán ordinarias las que se celebren periódicamente, cuando menos una vez al mes, conforme al calendario estipulado en el plan de trabajo.
- b) Serán extraordinarias las que se celebren en caso necesario y previa solicitud de la mayoría de los asesores técnicos, o bien, a petición del Secretario Técnico.

**XVI.** Que el artículo 366, numeral 1, del Reglamento, establece lo siguiente:

- Cada Comité Técnico Asesor de los Conteos Rápidos, a través de su Secretario Técnico, podrá convocar a sus sesiones a funcionarios de la autoridad electoral correspondiente y, en su caso, a especialistas cuyos conocimientos y experiencia aporten elementos relevantes para el desahogo de aspectos o temas específicos, quienes asistirán en calidad de invitados, con derecho a voz.
- Los Consejeros Electorales del Consejo General del Instituto Nacional Electoral o del Organismo Público Local correspondiente, podrán asistir a las sesiones o designar un representante, y tendrán derecho a voz.

**XVII.** Que el artículo 367, del Reglamento, menciona que el Comité Técnico Asesor de los Conteos Rápidos, tendrá las funciones siguientes:

- a) Proponer el plan de trabajo y el calendario de sesiones;
  - b) Proponer los criterios científicos, logísticos y operativos que se utilizarán en la estimación de los resultados de los conteos rápidos, y para normar el diseño y selección de la muestra;
  - c) Poner a consideración del Consejo General del Instituto Nacional Electoral o del Órgano Superior de Dirección del Organismo Público Local, según sea el caso, la aprobación de los criterios científicos, logísticos y operativos, mismos que deberán cumplir con lo previsto en el Capítulo III, del Título III, del mismo ordenamiento legal;
  - d) Coadyuvar con el Instituto Nacional Electoral o el Organismo Público Local, según corresponda, en la supervisión del cumplimiento del diseño, implementación y operación de los conteos rápidos;
  - e) Durante la jornada electoral, recibir la información de campo después del cierre de casillas; analizarla y realizar una estimación de los resultados de la elección. En caso de no poder realizar dicha estimación, deberán justificarlo;
  - f) Garantizar el uso responsable de la información estadística, a propósito de su función, atendiendo el procedimiento de resguardo de la muestra, a fin de dotar de confiabilidad a los conteos rápidos; y
  - g) Presentar los informes señalados en el artículo 368 del propio Reglamento.
- XVIII.** Que de conformidad con el artículo 368, numerales 1 y 3, del Reglamento, el Comité Técnico Asesor de los Conteos Rápidos, deberá presentar ante el Consejo General del Instituto Nacional Electoral u Órgano Superior de Dirección del Organismo Público Local, según corresponda, informes mensuales sobre los avances de sus actividades, durante los quince días posteriores a la fecha de corte del periodo que reporta; de igual forma al término de su encargo, deberá presentar al Consejo General del Instituto Nacional Electoral u Órgano Superior de Dirección que corresponda, un informe final de las actividades desempeñadas y de los resultados obtenidos en los conteos rápidos, así como las recomendaciones que consideren pertinentes, a más tardar cuarenta días después de la jornada electoral.
- XIX.** Que en términos del artículo 11, párrafos primero y décimo tercero, de la Constitución Política del Estado Libre y Soberano de México, la organización, desarrollo y vigilancia de los procesos electorales para las elecciones de Gobernador, Diputados a la Legislatura del Estado y miembros de Ayuntamientos es una función que se realiza a través del Instituto Nacional Electoral y el Organismo Público Electoral del Estado de

México, denominado Instituto Electoral del Estado de México, y que tendrá a su cargo, además de las que determine la ley de la materia, las actividades relativas a los resultados preliminares y conteo rápido, entre otras.

- XX.** Que como lo establece el artículo 175, del Código, el Consejo General es el Órgano Superior de Dirección del Instituto Electoral del Estado de México, responsable de vigilar el cumplimiento de las disposiciones constitucionales y legales en materia electoral, así como de velar porque los principios de certeza, legalidad, independencia, imparcialidad, máxima publicidad, objetividad y profesionalismo, guíen todas las actividades del Instituto.
- XXI.** Que en términos de lo previsto por los artículos 183, párrafos primero, segundo y cuarto, así como su fracción II, del Código, 1.3, párrafo primero y fracción II, último párrafo, así como el 1.4 y 1.6, párrafos primero y tercero del Reglamento para el Funcionamiento de las Comisiones del Consejo General del Instituto Electoral del Estado de México, destaca lo siguiente:
- El Consejo General integrará las comisiones que considere necesarias para el desempeño de sus atribuciones.
  - Las comisiones serán integradas, por tres consejeros designados por el Consejo General con voz y voto, por los representantes de los partidos con voz y un secretario técnico que será designado por el Consejo General en función de la Comisión de que se trate.
  - La aprobación de todos los acuerdos y dictámenes deberá ser con el voto de al menos dos de los integrantes, y preferentemente con el consenso de los partidos y coaliciones.
  - Bajo ninguna circunstancia las circulares, proyectos de acuerdo o de dictamen que emitan, tendrá obligatoriedad, salvo el caso de que sean aprobadas por el Consejo General. Solo en este supuesto podrán ser publicados en el periódico oficial "Gaceta del Gobierno".
  - Las comisiones especiales serán aquellas que se conformarán para la atención de las actividades sustantivas del Instituto que por su especial naturaleza, no tienen el carácter de permanente.
- XXII.** Que en términos del artículo 185, fracción I, del Código, es atribución de este Consejo General, expedir entre otros aspectos, las disposiciones que sean necesarias para el buen funcionamiento del Instituto.
- XXIII.** Que como se ha referido en el Resultado 4 del presente Acuerdo, el motivo de creación de la Comisión Especial para la atención del Programa

de Resultados Electorales Preliminares y Conteos Rápidos, fue a este Consejo General en la coordinación de las actividades relativas a la realización de los conteos rápidos, en la elección de Titular del Poder Ejecutivo del Estado de México, entre otros aspectos.

Ahora bien, tomando en consideración que para el presente Proceso Electoral en el que se elegirá Gobernador Constitucional del Estado de México, este Instituto Electoral deberá realizar conteos rápidos, como lo dispone el artículo 357, numeral 2, del Reglamento de Elecciones del Instituto Nacional Electoral.

#### **IV. OBJETIVO DEL COTECORA**

##### **General.**

Ofrecer los elementos necesarios para el diseño, implementación y operación de los conteos rápidos, como mecanismo para difundir estimaciones de resultados de las elecciones de gobernador a celebrarse el 4 de junio de 2017.

##### **Específicos**

- a) Elaborar y poner a consideración del Consejo General o de la instancia correspondiente, los criterios científicos, logísticos y operativos que se utilizarán en la estimación de los resultados en los Conteos Rápidos, y para normar el diseño y selección de la muestra;
- b) Coadyuvar con el Instituto Electoral del Distrito Federal, en la supervisión del cumplimiento del diseño, implementación y operación de los Conteos Rápidos.
- c) Durante la jornada electoral recibir la información de campo después del cierre de casillas, analizarla y determinar una estimación de los resultados de la elección.
- d) Garantizar el uso responsable de la información estadística que dispongan, a propósito de su función, haciendo plena observancia del resguardo de la muestra para garantizar la confiabilidad del conteo rápido.

#### **V. OBJETIVO DE LOS CONTEOS RÁPIDOS.**

##### **General.**

El objetivo general de los conteos rápidos es proporcionar certidumbre, confianza y transparencia en el proceso electoral, ofreciendo una estimación de alta calidad estadística de las tendencias de la votación el mismo día de la elección.

### **Específico**

El objetivo del Conteo Rápido en la elección de Gobernador en el estado de México será estimar el porcentaje de votos a favor de los candidatos a gobernador y el porcentaje de participación en la elección.

## **VI. PROGRAMA DE ACTIVIDADES**

El programa de trabajo del COTECORA consistió en tres líneas de trabajo mismas que atendieron el objetivo general y los objetivos específicos.

La primera línea de trabajo es referente a las actividades del COTECORA para definir los elementos técnicos y metodológicos que consideró el ejercicio de los conteos rápidos, así como los criterios de difusión de las tendencias electorales el día de la jornada electoral. En este sentido, las actividades implementadas fueron:

- i. Definir los elementos técnicos y metodológicos que consideró el mecanismo usado para estimar las tendencias de los resultados de la elección a Gobernador. Entre los elementos considerados están la población objetivo, procedimiento de estimación, tamaño y forma de obtención de la muestra, calidad de la estimación, la unidad de observación, marco muestral, la unidad de muestreo y otras variables de interés.
- ii. Proponer al Consejo General del IEEM los procedimientos de resguardo de la muestra original y sus copias, así como los períodos que ampararon su custodia.
- iii. Asesorar en la instrumentación del sistema informático que se utilizó para capturar los resultados de las actas de escrutinio y cómputo de las casillas de la muestra.
- iv. Realizar la programación y ejecución de las fórmulas estadísticas y algoritmos matemáticos que se utilizaron para procesar y analizar los resultados de la elección y, en consecuencia, la estimación de las tendencias electorales.
- v. Dar seguimiento al desarrollo del sistema informático.

- vi. Definir los criterios de difusión de la información el mismo día de la jornada electoral y presentarlos ante la instancia correspondiente.
- vii. Realizar la selección de la muestra a utilizar en el ejercicio de los conteos rápidos en el acto público convocado para tal fin.
- viii. Durante la jornada electoral, recibir la información de campo después del cierre de casillas, analizarla y realizar una estimación de los resultados de la elección de gobernador.
- ix. Presentar el reporte de resultados y sus conclusiones la noche de la elección al Consejo General del IEEM.
- x. Garantizar el uso responsable de la información estadística, a fin de dar confiabilidad al proceso de conteo rápido.

Una segunda línea de trabajo del COTECORA se relacionó con la asesoría en la implementación del operativo en campo utilizado para recabar y transmitir los resultados de la votación asentados en las actas de escrutinio y cómputo de las casillas seleccionadas en la muestra.

En este contexto, las actividades involucradas fueron las siguientes:

- i. Coadyuvar con el área responsable para definir las características del operativo en campo que se utilice en los conteos rápidos.
- ii. Brindar apoyo en la confección de los criterios para el proceso de selección del personal que se contrate para el ejercicio de conteos rápidos.
- iii. Opinar y, en su caso, aportar elementos de mejora para los esquemas de capacitación del personal involucrado en el operativo en campo.
- iv. Sugerir estrategias y acciones a implementar ante situaciones no previstas en la recolección y transmisión de los datos.
- v. Opinar y, en su caso, aportar elementos para la definición de la logística que se empleará en el operativo en campo.
- vi. Supervisar y dar seguimiento a la prueba de captura y a los simulacros de conteos rápidos que ejecuten las instancias responsables.

La tercera de ellas consiste en atender las actividades que dan sustento al funcionamiento del COTECORA. Entre estas actividades se encuentran:

- i. Instalar el COTECORA.

- ii. Aprobar su Plan de Trabajo.
- iii. Aprobar su Calendario de Sesiones Ordinarias.
- iv. Aprobar su Manual de Funcionamiento del COTECORA.
- v. Establecer los criterios científicos, logísticos y operativos que se utilizarían en la estimación del resultado de la elección a gobernador en 2017, y ponerlos a consideración del Consejo General.
- vi. Proponer al Consejo General el protocolo de selección de la muestra, los procedimientos de resguardo de la misma y los períodos de su custodia.
- vii. Elaborar y aprobar los informes mensuales de actividades del COTECORA y remitirlos al Consejo General del Instituto.
- viii. En su caso, aprobar el informe final sobre la implementación de los mecanismos para estimar los resultados electorales el día de la jornada electoral.
- ix. Elaborar y aprobar el informe final de actividades del COTECORA.

## **VII. SESIONES DEL COTECORA**

El COTECORA sesionó en forma ordinaria:

En 2016	3 de diciembre
En 2017	25 de enero
	22 de febrero
	22 de marzo
	26 de abril
	24 de mayo

Reunión preparatoria      16 de diciembre 2016

La forma de trabajar del COTECORA motivó a que hubiera diversas sesiones de trabajo:

Reuniones de trabajo:

11 de enero
18 de enero
1 de marzo
3 de abril
19 de abril
3 de mayo
8 de mayo

Asimismo, miembros de COTECORA asistieron a las sesiones de la Comisión Especial para la Atención del Programa de Resultados electorales Preliminares y Conteos Rápidos el 24 de febrero y el 9 de mayo.

Existieron además reuniones de trabajo con Consejeros del IEEM y Consejeros del INE. Éstas se celebraron el 21 y 27 de abril.

Por otro lado, miembros del COTECORA asistieron a TV mexicana el 24 de abril a dialogar sobre los Conteos Rápidos a implementarse en la elección de Gobernador en el estado de México.

Miembros del COTECORA asistieron a la capacitación de los vocales para la utilización de los sistemas de cómputo. Esto fue realizado el 9 de mayo.

Reuniones de trabajo después de los simulacros con partidos y consejeros:

- 14 de mayo
- 21 de mayo
- 28 de mayo

## **VIII. CRITERIOS CIENTÍFICOS**

Los criterios científicos del Conteo Rápido de la Elección del Estado de México son todos los procedimientos que, con base en la probabilidad y en la estadística, se usan para estimar el porcentaje de votos a favor de los candidatos para gobernador del estado.

### **1. Motivación**

El objetivo general de los conteos rápidos es crear certidumbre, confianza y transparencia en el proceso electoral, ofreciendo una estimación de las tendencias de la votación de alta calidad estadística en la noche del mismo día de la elección.

Los objetivos particulares del Conteo Rápido para la Elección de Gobernador en el Estado de México fueron:

- Estimar el porcentaje de votos en favor de cada aspirante a gobernador por el Estado de México.
- Estimar el porcentaje de participación en la elección.

Todas las estimaciones del COTECORA fueron diseñadas para ser generadas en forma de intervalos de confianza del 95%.

## 2. Esquema de muestreo

### Diseño de muestreo

Las estrategias de estimación diseñadas por los miembros del COTECORA fueron basadas en la misma información muestral. Se consideró el proceso de selección a través de un muestreo aleatorio estratificado, en donde al interior de cada estrato se seleccionarían casillas mediante un muestreo aleatorio simple sin reemplazo.

Para definir los estratos, se usaron los 45 Distritos Locales del Estado de México, (ver Tabla 1). Primeramente, se consideró una muestra total de alrededor de 2,000 casillas, esperando recibir al menos 1,200 casillas para realizar la estimación final. Vale la pena resaltar que el tamaño de muestra definitivo se calculó con base a la lista final de casillas a instalarse el 4 de junio y considerando la tasa de no respuesta de cada distrito local observada en elecciones pasadas.

La distribución de casillas en cada estrato se realizó vía afijación proporcional.

### Consideraciones para definir el tamaño de muestra y diseño de muestreo

Para determinar el tamaño de muestra en para el Conteo Rápido del Estado de México fue probar distintas estratificaciones y tamaños de muestra eligiendo la combinación que permitiera obtener la mejor precisión, pero cuidando no rebasar la capacidad operativa de campo.

En todos los ejercicios realizados se tomaron como referencia los cómputos distritales de tres votaciones:

1. Elección para Gobernador en el Estado de México del 3 de julio de 2011.
  - 17,489 casillas en total  $\Rightarrow$  17,472 casillas para análisis (17 casillas).
2. Elección Presidencial, tomando en cuenta sólo los resultados en el Estado de México, del 1 de julio de 2012.
  - 17,350 casillas en total  $\Rightarrow$  17,229 casillas para análisis (121 casillas).
3. Elección Federal de Diputados del 7 de junio de 2015. En este caso se asumió que los candidatos fueron los mismos en los 40 Distritos Federales y que hubo dos coaliciones: PRI-PVEM y PRD-PT.

- 18,250 casillas en total  $\Rightarrow$  18,115 casillas para análisis (135 casillas).

La diferencia entre las casillas totales y las casillas para análisis se debió a que no fue posible determinar el DISTRITO LOCAL de algunas casillas. Sin embargo, se consideró que esto no alteraba el fondo del análisis para determinar el tamaño de muestra.

### Estratificaciones y tamaños de muestra

Las estratificaciones consideradas en los ejercicios de simulación, para cada elección, fueron seis:

1. Distritos Federales  $\Rightarrow$  40 estratos.
2. Distritos Federales y al interior de cada distrito se dividió por sección urbana no urbana  $\Rightarrow$  64 estratos.
3. Distritos Locales  $\Rightarrow$  45 estratos.
4. Distritos Locales y al interior de cada distrito se dividió por sección urbana y no urbana  $\Rightarrow$  75 estratos.
5. Municipios  $\Rightarrow$  125 estratos.
6. Municipios y al interior de cada municipio se dividió por sección urbana y no urbana  $\Rightarrow$  206 estratos.

Para distribuir el tamaño de muestra en cada uno de los estratos, se utilizó lo que se conoce como “afijación proporcional”:

$$n_h = n \frac{N_h}{N},$$

en donde  $n_h$  es el tamaño de muestra que se seleccionará en el estrato  $h$ ,  $n$  el tamaño de muestra total,  $N_h$  el número total de casillas instaladas en el estrato  $h$  y  $N$  el número total de casillas instaladas<sup>1</sup> en el Estado de México. En palabras comunes, el tamaño de muestra total, se distribuirá proporcionalmente de acuerdo al número de casillas instaladas en cada estrato. En estratos con mayor número de casillas se seleccionará mayor muestra.

### Resultados de las simulaciones

Lo que se busca es la estratificación que alcance la mejor precisión (o error máximo aceptable), con el menor tamaño de muestra posible. Experiencias anteriores indican que una precisión de

$$\epsilon = 0.5(\text{medio punto porcentual}).$$

---

<sup>1</sup>Se toman las casillas para análisis descritas en la sección anterior.

es una buena alternativa y aunque exige tamaños de muestra grandes, no son prohibitivos y son manejables en campo.

En todas las simulaciones, se utilizó el estimador de razón dado por la expresión (3), los resultados obtenidos con cada una de las tres Elecciones, se muestran en las Tablas 2, 3 y 4 (para obtener las precisiones en cada celda de la tabla se seleccionaron 50,000 muestras independientes y con cada muestra se hizo una estimación, ver Gráfica 1).

Tabla 2. Elección para Gobernador 2011: precisiones obtenidas usando el estimador (3), para distintos tamaños de muestra y estratificaciones.

TAMAÑO DE MUESTRA	DISTRITOS FEDERALES	DISTRITOS FEDERALES URBANO - NO URBANO	DISTRITOS LOCALES	DISTRITOS LOCALES URBANO - NO URBANO	MUNICIPIOS	MUNICIPIOS URBANO - NO URBANO
700	0.62	0.62	0.60	0.60	0.58	0.56
750	0.60	0.59	0.58	0.57	0.56	0.54
800	0.58	0.57	0.56	0.55	0.55	0.53
850	0.56	0.55	0.55	0.53	0.53	0.52
900	0.54	0.53	0.53	0.52	0.51	0.50
950	0.53	0.53	0.51	0.51	0.49	0.48
1,000	0.52	0.51	0.50	0.49	0.48	0.47
1,050	0.50	0.50	0.48	0.49	0.47	0.46
1,100	0.49	0.48	0.47	0.47	0.46	0.45
1,150	0.48	0.48	0.46	0.46	0.45	0.44
1,200	0.47	0.47	0.45	0.45	0.44	0.43
1,250	0.46	0.45	0.44	0.43	0.43	0.42
1,300	0.45	0.44	0.43	0.43	0.42	0.41
1,350	0.44	0.44	0.42	0.42	0.42	0.41
1,400	0.44	0.43	0.42	0.41	0.40	0.39

Tabla 3. Elección Presidencial 2012: precisiones obtenidas usando el estimador (3), para distintos tamaños de muestra y estratificaciones.

TAMAÑO DE MUESTRA	DISTRITOS FEDERALES	DISTRITOS FEDERALES URBANO - NO URBANO	DISTRITOS LOCALES	DISTRITOS LOCALES URBANO - NO URBANO	MUNICIPIOS	MUNICIPIOS URBANO - NO URBANO
700	0.55	0.54	0.52	0.52	0.51	ALGUNOS ESTRATOS CON 1 ÚNICA CASILLA INSTALADA
750	0.53	0.52	0.50	0.50	0.49	
800	0.51	0.50	0.49	0.48	0.48	
850	0.49	0.48	0.48	0.46	0.46	
900	0.48	0.47	0.46	0.45	0.45	
950	0.46	0.46	0.44	0.44	0.43	
1,000	0.45	0.45	0.44	0.43	0.42	

<b>1,050</b>	0.44	0.43	0.42	0.42	0.42	
<b>1,100</b>	0.43	0.42	0.41	0.41	0.40	
<b>1,150</b>	0.42	0.41	0.40	0.40	0.40	
<b>1,200</b>	0.41	0.40	0.39	0.39	0.39	
<b>1,250</b>	0.40	0.39	0.38	0.38	0.38	
<b>1,300</b>	0.39	0.39	0.38	0.37	0.37	
<b>1,350</b>	0.38	0.38	0.37	0.36	0.36	
<b>1,400</b>	0.37	0.37	0.36	0.36	0.35	

Tabla 4. Elección Federal de Diputados 2015: precisiones obtenidas usando el estimador (3), para distintos tamaños de muestra y estratificaciones.

TAMAÑO DE MUESTRA	DISTRITOS FEDERALES	DISTRITOS FEDERALES URBANO - NO URBANO	DISTRITOS LOCALES	DISTRITOS LOCALES URBANO - NO URBANO	MUNICIPIOS	MUNICIPIOS URBANO - NO URBANO
700	0.68	0.67	0.66	0.66	0.59	0.56
750	0.67	0.65	0.64	0.63	0.57	0.55
800	0.64	0.62	0.61	0.61	0.55	0.53
850	0.62	0.61	0.59	0.59	0.52	0.51
900	0.60	0.59	0.58	0.57	0.51	0.50
950	0.58	0.58	0.56	0.56	0.50	0.48
1,000	0.57	0.55	0.55	0.54	0.49	0.48
1,050	0.55	0.54	0.54	0.52	0.48	0.47
1,100	0.53	0.53	0.52	0.51	0.47	0.45
1,150	0.52	0.52	0.51	0.50	0.45	0.44
1,200	0.51	0.51	0.50	0.49	0.44	0.44
1,250	0.51	0.49	0.48	0.47	0.43	0.42
1,300	0.50	0.48	0.48	0.47	0.43	0.42
1,350	0.48	0.48	0.46	0.46	0.42	0.40
1,400	0.47	0.47	0.46	0.45	0.41	0.40

Estos resultados mostraron que, de las tres elecciones, la que requeriría el menor tamaño de muestra para alcanzar una mejor precisión sería la Presidencial, seguida por la Elección de Gobernador y la que exigiría un mayor tamaño de muestra es la Elección de Diputados. Esto es debido a que la varianza de la votación es mayor en la Elección de Diputados, en nuestro análisis supusimos que hubo 8 candidatos (con 2 coaliciones) en los 40 Distritos Federales que componen el Estado de México, sin embargo, en la realidad en cada distrito hubo candidatos distintos. Esto introduce una gran variabilidad en la votación para todo el estado. Así pues, se consideró que el tamaño de muestra que se derive del análisis de la votación para Diputados necesariamente sería suficiente para las otras votaciones.

En la Tabla 3, la mejor estratificación es la que contempla a los municipios, ya que el tamaño de muestra requerido para alcanzar un error de  $\epsilon = 0.5$  es de  $n = 950$  casillas. Sin embargo, existe un problema práctico con esta estratificación. El detalle se resume en la Tabla 4.

Tabla 4. Elección Federal de Diputados 2015: Estadísticas básicas para cada estratificación y las casillas instaladas en cada estrato.

Variable de interés	DISTRITOS FEDERALES	DISTRITOS FEDERALES URBANO - NO URBANO	DISTRITOS LOCALES	DISTRITOS LOCALES URBANO - NO URBANO	MUNICIPIOS	MUNICIPIOS URBANO - NO URBANO
Número de estratos	40	64	45	75	125	194
Promedio de casillas instaladas	453	283	403	242	145	93
Máximo casillas instaladas	678	606	522	522	1,966	1,966
Mínimo casillas instaladas	350	3	302	2	6	2

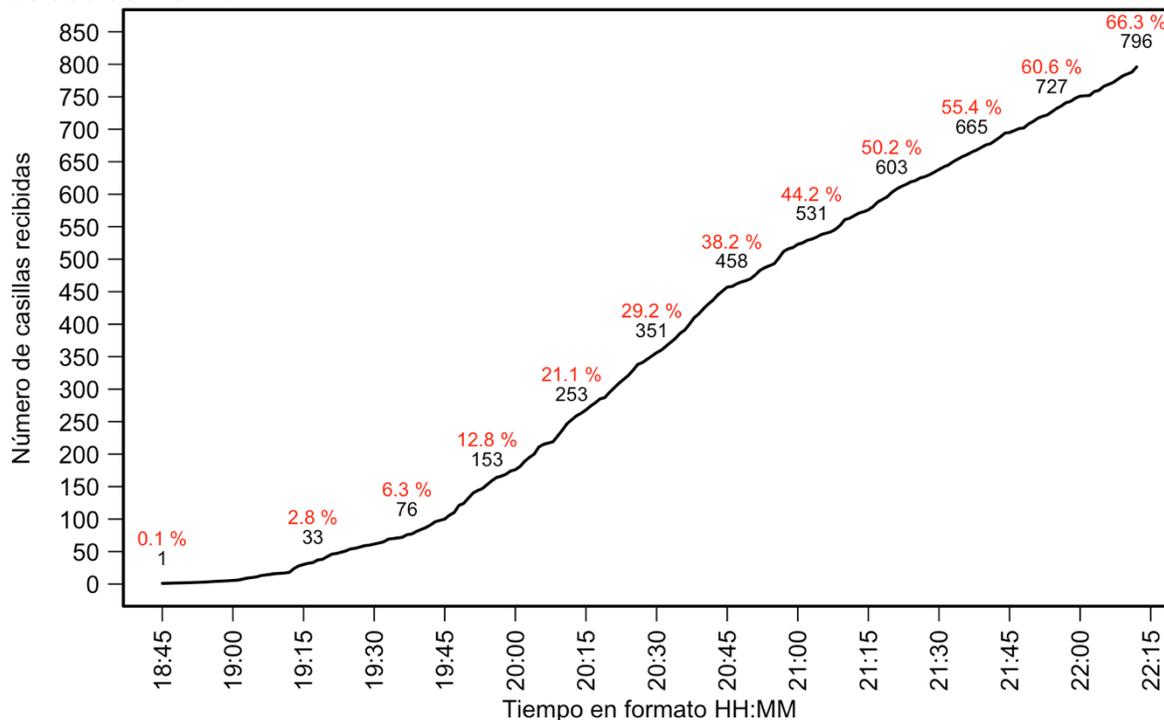
El mínimo de casillas instaladas en uno de los municipios es 6, de hecho, en 23 municipios se instalaron menos de 20 casillas. Esto representa un gran riesgo para el conteo rápido, de elegirse los municipios como estratos, de cada municipio se extraería una muestra de casillas y con la información que se recupere de esta muestra, se realiza una estimación. Finalmente, se conjuntan las estimaciones de todos los estratos para obtener la estimación global, ver ecuación (3). Con estratos con tan pocas casillas instaladas, aunque se extraiga una muestra grande (proporcionalmente), la probabilidad de que no se recupere ninguna casilla en estos estratos es muy alta. Si no se tiene información en algún estrato, será imposible realizar la estimación. Lo mismo ocurre si se utiliza alguna estratificación que contemple la parte urbana y no urbana, el número de casillas instaladas en algunos estratos es muy pequeño. Este riesgo se reduce de manera sustancial, si se toman los Distrito Locales como estratos, en este caso el mínimo de casillas instaladas es de 302, el máximo es de 522 y en promedio se instalan 403 casillas por Distrito Local. Además, también se alcanza una buena precisión con un tamaño de muestra aceptable.

Tomando los Distritos Locales como estratos, se obtendría una precisión de  $\epsilon = 0.5$  con  $n = 1,200$  casillas, además, utilizando la afijación proporcional el promedio de casillas en muestra por estrato sería de 27, con un máximo de 35 y mínimo de 20.

#### *No respuesta o casillas no reportadas*

En cualquier encuesta existe un porcentaje de no respuesta, en el caso del Conteo Rápido serían casillas seleccionadas en muestra, pero que por diversos motivos su información no se recibe a tiempo para hacer la estimación. En el Conteo Rápido para la Elección Federal de Diputados de 2015, la muestra en el Estado de México fue de 1,200 casillas (30 casillas por cada Distrito Federal), sin embargo, a las 10:00 pm sólo se había recibido el 60 % y a las 10:15 pm, cuando se hizo la última estimación, se tenía aproximadamente el 66 % de la muestra total. El número de casillas recibidas, así como la hora de arribo se muestra en la Gráfica 2.

**Gráfica 2.** Conteo Rápido 2015: Serie de tiempo para el número de casillas recibidas del Estado de México.



Así pues, en un primer ejercicio se consideró que si el mismo comportamiento se mantiene para el Conteo Rápido en la Elección de Gobernador en 2017, para realizar la estimación final a las 10 pm se necesitaría un tamaño de muestra total, considerando la no respuesta, de

$$n_{Total} = \frac{1,200}{0.6} = 2,000 \text{ casillas.}$$

Con este tamaño de muestra total, se esperaría contar con un tamaño de muestra efectivo de  $n = 1,200$  casillas alrededor de las 10 pm del 4 de junio. Posteriormente, sin embargo, con información recibida por parte de la Junta Local ejecutiva del INE en el Estado de México, se consideró posible reducir el porcentaje de no respuesta o de respuesta a destiempo a lo observado a las 22:15 en la gráfica anterior. Así pues, se concluyó que la muestra debería de ser de:

$$n_{Total} = \frac{1,200}{0.66} = 1,818 \text{ casillas.}$$

Al momento del diseño es importante tomar en cuenta los siguientes puntos:

1. La precisión se maneja tanto en la etapa de planeación del ejercicio, como en la presentación de los resultados finales del mismo. En la etapa de

planeación, se usa para determinar un tamaño de muestra capaz (teóricamente) de alcanzar el nivel de precisión y confianza deseadas en la estimación. Sin embargo, debido a que este tamaño de muestra se fija tomando como referencia los resultados de una o varias elecciones previas, una vez concluida la elección y con toda la información disponible se determina la precisión realmente obtenida.

2. Los resultados de las simulaciones presentadas son únicamente para el método de estimación usado (ecuación 3), y tienen como objetivo fijar un tamaño de muestra. Sin embargo, se pueden emplear métodos de estimación alternativos (usando la misma estrategia de selección) que arrojarán precisiones distintas.
3. Para la elección en el Estado de México, se esperaba contar con 3,352 capacitadores asistentes electorales (CAES) que recabarán la información de las casillas seleccionadas en muestra y procederán a su transmisión. Es conveniente apuntar que el tamaño de muestra se encontraba dentro de la capacidad operativa de campo.

### **DISTRIBUCIÓN DE LA MUESTRA POR ESTRATO**

La distribución de la muestra en cada uno de los distritos locales, fue la siguiente:

<b>DSITRIBUCIÓN POR ESTRATO DE LAS 1818 CASILLAS DE LA MUESTRA</b>				
<b>DISTRITO_LOC</b>	<b>LN</b>	<b>Ncasillas</b>	<b>ESTRATO</b>	<b>NcasillasE</b>
1	245,295	370	1	36
2	228,852	358	2	35
3	241,066	365	3	36
4	246,484	393	4	38
5	249,813	389	5	38
6	213,716	341	6	33
7	248,510	418	7	41
8	225,127	382	8	37
9	261,131	521	9	51
10	238,123	433	10	42
11	236,537	382	11	37
12	245,386	386	12	38
13	248,809	439	13	43
14	241,414	446	14	44
15	238,904	397	15	39
16	251,088	391	16	38
17	237,368	378	17	37
18	285,900	496	18	48
19	254,349	386	19	38
20	291,938	462	20	45
21	236,380	387	21	38
22	220,256	375	22	37
23	271,806	424	23	41
24	271,108	470	24	46
25	289,109	515	25	50
26	256,333	429	26	42
27	251,454	394	27	39
28	242,018	398	28	39
29	270,936	461	29	45
30	257,598	444	30	43
31	246,782	392	31	38
32	263,363	418	32	41
33	304,726	474	33	46
34	262,679	433	34	42
35	249,302	393	35	38
36	239,715	364	36	36
37	244,139	379	37	37
38	260,304	415	38	41
39	258,304	427	39	42
40	246,248	411	40	40
41	282,493	500	41	49
42	211,926	370	42	36
43	252,537	419	43	41
44	269,486	416	44	41
45	224,105	364	45	36

## Estimación e integración de los intervalos de confianza

Para dar mayor certeza a la estimación, los integrantes del COTECORA utilizaron tres enfoques estadísticos distintos para estimar los resultados de la elección:

- Método Bayesiano, a cargo del Dr. Arturo Erdely Ruiz.
- Método Clásico vía las fórmulas para estimadores de razón, derivadas mediante las técnicas de muestreo probabilístico, a cargo de la Dra. Karla Beatriz Valenzuela Ocaña y el Dr. Alberto Alonso y Coria.
- Método Clásico vía métodos de remuestreo, a cargo del Dr. Carlos E. Rodríguez.

Dado que se aplicaron diversos métodos de estimación, los límites inferior y superior de cada intervalo no fueron estrictamente idénticos, toda vez que los métodos que se utilizarán introducen elementos aleatorios. Fue por tanto necesario establecer un criterio para conjuntar los resultados.

Las estrategias de estimación que implementaron los miembros del COTECORA fueron basadas en la misma muestra. Durante toda la jornada, el COTECORA monitoreó las tendencias de votación indicadas por los intervalos de confianza.

Para integrar las estimaciones en una sola, se aplicó el siguiente procedimiento: Sea  $I_k = [a_k, b_k]$  una estimación por intervalo de un porcentaje de votación, donde el subíndice  $k = 1, \dots, r$  representa cada una de las  $r$  estimaciones que realizarán los miembros del COTECORA. El intervalo combinado  $I = [a, b]$  que reportará el COTECORA se obtendrá mediante las fórmulas:

$$a = \frac{\text{promedio}\{a_1, \dots, a_r\} + \text{mediana}\{a_1, \dots, a_r\}}{2} \quad \text{y} \quad b = \frac{\text{promedio}\{b_1, \dots, b_r\} + \text{mediana}\{b_1, \dots, b_r\}}{2}$$

Utilizar solo promedios tiene la ventaja de incorporar directamente información de todos los intervalos, pero la desventaja de que un solo valor atípico podría tener demasiada influencia. Utilizar solo medianas tiene la ventaja de ser inmune ante un valor atípico, pero la desventaja de que varios intervalos podrían quedar totalmente fuera de consideración. Es por lo anterior que se escogió un término medio entre ambas posibilidades.

## Procedimientos de estimación

### Propuesta del Dr. Arturo Erdely Ruiz

# Estimación bayesiana por intervalos de porcentajes de votación

Dr. Arturo Erdely Ruiz

Abril de 2017

## Resumen

Se propone la utilización de la distribución de probabilidad multinomial, y la correspondiente inferencia bayesiana de sus parámetros, para la estimación por intervalo de los porcentajes del voto emitido que obtendrán los candidatos en la elección para gobernador del Estado de México del año 2017, a partir de un conteo rápido de actas electorales por casilla.

## 1. Introducción

Si una variable o vector aleatorio  $X$  con una función de densidad o masa de probabilidades  $f(x|\theta)$  cuantifica adecuadamente la incertidumbre sobre un fenómeno aleatorio de interés, pero su parámetro (o vector de parámetros)  $\theta \in \mathfrak{S}$  es desconocido, donde  $\mathfrak{S}$  es el conjunto de valores admisibles para  $\theta$  (espacio paramétrico), el enfoque bayesiano de la estadística propone cuantificar la incertidumbre sobre  $\theta$  mediante una variable aleatoria  $\Theta$  con una distribución de probabilidad sobre  $\mathfrak{S}$ , misma que puede basarse en información inicial con que se cuente en un momento dado y que se denomina *distribución a priori*, y que al combinarse con información muestral se convierte en una *distribución a posteriori*.

Sea  $(X_1, \dots, X_m)$  una muestra aleatoria con función de densidad o masa de probabilidades conjunta  $f(y_1, \dots, y_m|\theta)$ . Si  $(x_1, \dots, x_m)$  son los valores observados de dicha muestra, y se denota mediante  $\pi(\theta)$  a la función de densidad de probabilidades a priori, y mediante  $p(\theta|x_1, \dots, x_m)$  a la función de densidad de probabilidades a posteriori, entonces mediante la *Regla de Bayes* se obtiene la siguiente fórmula:

$$p(\theta|x_1, \dots, x_m) = \frac{1}{B} \pi(\theta) f(x_1, \dots, x_m|\theta), \quad \theta \in \mathfrak{S}, \quad (1)$$

donde la constante  $B = \int_{\mathfrak{S}} \pi(\theta) f(x_1, \dots, x_m|\theta) d\theta$ . Mediante (1) se realiza cualquier inferencia estadística de interés sobre el parámetro (o vector de parámetros)  $\theta$ . Para el caso de *estimación puntual*, misma que consiste en escoger un valor  $\hat{\theta} \in \mathfrak{S}$  con base en toda la información disponible, puede calcularse alguna medida de tendencia central (la media o la mediana, por ejemplo) por medio de  $p(\theta|x_1, \dots, x_m)$ .

En el caso de *estimación por intervalo* de probabilidad  $0 < \alpha < 1$ , si  $\theta$  es de dimensión 1 entonces se busca un intervalo  $I(\alpha) := [a, b] \subset \mathfrak{S}$  tal que:

$$\mathbb{P}[\theta \in I(\alpha)] = \int_a^b p(\theta|x_1, \dots, x_m) d\theta = \alpha. \quad (2)$$

Usualmente la solución para  $I(\alpha)$  en (2) no es única, por lo que de entre el conjunto de soluciones posibles  $\{[a, b] : \int_a^b p(\theta | x_1, \dots, x_m) d\theta = \alpha\}$  se escoge aquel intervalo cuya longitud  $b - a$  sea mínima. En caso de que la dimensión de  $\mathfrak{S}$  sea  $J > 1$  entonces la densidad de probabilidades a posteriori (1) de  $\theta = (\theta_1, \dots, \theta_J)$  es multivariada y por tanto deberá marginalizarse para cada componente  $\theta_j$ , y mediante  $p(\theta_j | x_1, \dots, x_m)$  hacer estimación puntual y por intervalo para cada  $\theta_j$  con  $j \in \{1, \dots, J\}$ . Si  $p(\theta_j | x_1, \dots, x_m)$  no es simétrica respecto a algún valor  $\theta_j = \theta_0$ , obtener la estimación por intervalo  $I_j(\alpha)$  que sea de longitud mínima pudiera resultar un tanto laborioso o computacionalmente intensivo, y por tanto es común ocupar la solución pseudo-óptima que consiste en  $I_j(\alpha) = [a_j, b_j]$  tal que:

$$\int_{-\infty}^{a_j} p(\theta_j | x_1, \dots, x_m) d\theta_j = \frac{1 - \alpha}{2} = \int_{b_j}^{+\infty} p(\theta_j | x_1, \dots, x_m) d\theta_j. \quad (3)$$

## 2. Notación

Se adopta prácticamente tal cual la notación utilizada en Mendoza y Nieto-Barajas (2016). El *marco muestral* disponible consiste en un total de  $K$  casillas electorales que se instalarán el día de la elección para que un total de  $n$  posibles votantes (tamaño de la *lista nominal*) estén en condiciones de ejercer su derecho al voto, mismos que pueden decidir sufragar a favor de algún candidato registrado, o a favor de alguno no registrado (NOREG), o anular su voto (NULOS) o simplemente abstenerse de acudir a votar (ABST). Al conjunto de posibilidades anteriores les asignaremos las categorías  $1, 2, \dots, J$  reservando siempre  $J - 2$  para NOREG,  $J - 1$  para NULOS y  $J$  para ABST, esto es, los candidatos registrados serán  $1, \dots, J - 3$ , con  $J \geq 4$ . En lo sucesivo cuando se haga referencia a un *candidato*  $j \in \{1, \dots, J\}$  deberá entenderse en un sentido amplio, es decir considerando como “candidatos” incluso a las categorías NOREG, NULOS y ABST.

El total de  $K$  casillas electorales se reparten en  $N$  *estratos* o subconjuntos (distritos locales en este caso), y se denotará  $K_i$  al número de casillas en el distrito  $i \in \{1, \dots, N\}$  de modo que  $\sum_{i=1}^N K_i = K$ , y mediante  $n_i$  al tamaño de la lista nominal en el distrito  $i$  de modo que  $\sum_{i=1}^N n_i = n$ . Al momento del cierre de las casillas se procede al conteo de votos emitidos en cada una y su clasificación en las categorías  $1$  a la  $J - 1$ . Se representará mediante las variables aleatorias  $X_{i,j}^{(k)}$  el número de votos que obtendrá el candidato  $j \in \{1, \dots, J\}$  en el distrito  $i \in \{1, \dots, N\}$  y casilla  $k \in \{1, \dots, K_i\}$ . A su vez, las variables aleatorias  $X_{i,j} = \sum_{k=1}^{K_i} X_{i,j}^{(k)}$  representan el número total de votos que obtendrá el candidato  $j$  en el distrito  $i$ , y las variables aleatorias  $\Theta_{i,j} = X_{i,j}/n_i$  representan la proporción de votos a favor del candidato  $j$  en el distrito  $i$ . De lo anterior es inmediato que  $\sum_{j=1}^J X_{i,j} = n_i$  y  $\sum_{j=1}^J \Theta_{i,j} = 1$  para todo distrito  $i$ .

Con lo anterior es posible definir variables aleatorias  $X_j$  como el número total de votos que obtiene el candidato  $j$  y en consecuencia  $X_j = \sum_{i=1}^N X_{i,j} = \sum_{i=1}^N n_i \Theta_{i,j}$ . La proporción total de votos que obtendrá el candidato  $j$  se obtiene mediante las variables aleatorias:

$$\Theta_j = \frac{X_j}{n} = \sum_{i=1}^N \frac{n_i}{n} \Theta_{i,j}, \quad j \in \{1, \dots, J\} \quad (4)$$

de donde es inmediato que  $\sum_{j=1}^J X_j = n$  y en consecuencia  $\sum_{j=1}^J \Theta_j = 1$ . La proporción de votos  $\Theta_j$  es respecto a la lista nominal, pero también es de interés la proporción de votos respecto al total de

votos emitidos en las casillas (esto es, sin considerar la  $J$ -ésima categoría ABST), misma que puede representarse mediante las variables aleatorias:

$$\Lambda_j = \frac{X_j}{\sum_{\ell=1}^{J-1} X_\ell} = \frac{\Theta_j}{\sum_{\ell=1}^{J-1} \Theta_\ell}, \quad j \in \{1, \dots, J-1\}, \quad (5)$$

y la proporción de *participación* en la elección se calcula como complemento del abstencionismo (ABST), esto es como  $1 - \Theta_J$ .

Se realizará un *muestreo aleatorio estratificado* consistente en muestreo aleatorio simple de casillas por distrito (estrato) proporcional a la lista nominal de cada distrito. Esto es, si en el distrito  $i$  hay  $K_i$  casillas, de entre éstas se seleccionará al azar un subconjunto de  $c_i$  casillas donde  $c_i \ll K_i$  y de forma que si  $c = \sum_{i=1}^N c_i$  representa el número total de casillas en la muestra estratificada entonces  $c_i \approx \frac{n_i}{n} c$ .

### 3. Distribución de probabilidad multinomial

En virtud de que la *unidad de muestreo* para el conteo rápido es el resultado de la votación en cada casilla electoral, se propone representar los resultados de las distintas casillas mediante un vector aleatorio  $(X_{i,1}^{(k)}, \dots, X_{i,j}^{(k)})$  donde  $i \in \{1, \dots, N\}$  y  $k \in \{1, \dots, K_i\}$ . Los resultados que se asientan en el acta de cada casilla corresponden a votos efectivamente emitidos, esto es a los valores que se observarán para las variables aleatorias  $X_{i,1}^{(k)}, \dots, X_{i,J-1}^{(k)}$ , excluyendo la categoría  $J$  (ABSTenciones) ya que esta última se deduce restando al tamaño de la lista nominal de la casilla ( $n_{i,k}$ ) el total de la votación efectivamente emitida  $\sum_{j=1}^{J-1} X_{i,j}^{(k)}$  de modo que  $\sum_{j=1}^J X_{i,j}^{(k)} = n_{i,k}$ . Esto último implica una dependencia entre las variables aleatorias  $X_{i,j}^{(k)}$  ya que sin importar el resultado de cada casilla su suma siempre debe ser igual (teóricamente) al tamaño de la lista nominal de dicha casilla; así, por ejemplo, un valor muy elevado de alguna de ellas implica necesariamente valores modestos para las demás. Esto es, los valores de unas influyen en las demás porque el máximo número de votos obtenible por casilla ( $n_{i,k}$ ) ya está predefinido (salvo en unas pocas casillas denominadas *especiales*).

La idea de aplicar un muestreo aleatorio estratificado radica en el beneficio inferencial que se obtiene por establecer o escoger una estratificación tal que en cada estrato sea razonable suponer cierta homogeneidad en estas subpoblaciones, en contraste con la heterogeneidad que se tendría al obtener muestras directamente de la población total sin estratificar, en deterioro de la precisión de las inferencias que se desea realizar, véase por ejemplo Särndal *et al.* (1992). Con lo anterior, el vector aleatorio  $(X_{i,1}^{(k)}, \dots, X_{i,j}^{(k)})$  puede asociarse a un *experimento aleatorio multinomial* consistente en  $n_{i,k}$  repeticiones independientes (votantes en el distrito  $i$  y casilla  $k$ ) de un procedimiento que en cada repetición puede arrojar uno de entre  $J$  resultados posibles con una probabilidad  $0 < \theta_{i,j} < 1$ ,  $j \in \{1, \dots, J\}$ , con  $\sum_{j=1}^J \theta_{i,j} = 1$  (nótese que aquí se está aplicando el supuesto de homogeneidad por estrato, ya que las probabilidades teóricas  $\theta_{i,j}$  se suponen razonablemente homogéneas para toda casilla  $k$  dentro de un mismo distrito  $i$ ), y por tanto las variables aleatorias  $X_{i,j}^{(k)}$  representan el número de veces que se obtuvo el resultado  $j$  en las  $n_{i,k}$  repeticiones (número de votos a favor del candidato  $j$ ). Bajo los supuestos anteriores, la función de masa de probabilidades conjunta del vector aleatorio  $(X_{i,1}^{(k)}, \dots, X_{i,j}^{(k)})$  para cada distrito  $i$  y casilla  $k$ , véase por ejemplo Johnson *et al.* (1997),

sería **Multinomial**( $\theta_{i,1}, \dots, \theta_{i,J}, n_{i,k}$ ) :

$$\begin{aligned} f_i^{(k)}(z_1, \dots, z_J | \theta_{i,1}, \dots, \theta_{i,J}, n_{i,k}) &= \mathbb{P}(X_{i,1}^{(k)} = z_1, \dots, X_{i,J}^{(k)} = z_J) \\ &= \binom{n_{i,k}}{z_1, \dots, z_J} \prod_{j=1}^J \theta_{i,j}^{z_j} \mathbf{1}_{\{\sum_{j=1}^J z_{i,j} = n_{i,k}\}} \end{aligned} \quad (6)$$

Con lo anterior, el vector aleatorio de votos por candidato en cada distrito se obtiene sumando los votos obtenidos por cada candidato en cada casilla del distrito:

$$(X_{i,1}, \dots, X_{i,J}) = \left( \sum_{k=1}^{K_i} X_{i,1}^{(k)}, \dots, \sum_{k=1}^{K_i} X_{i,J}^{(k)} \right), \quad i \in \{1, \dots, N\} \quad (7)$$

donde de acuerdo a Johnson *et al.*(1997) el vector aleatorio (7) tiene distribución de probabilidad **Multinomial**( $\theta_{i,1}, \dots, \theta_{i,J}, n_i$ ) con  $n_i = \sum_{k=1}^{K_i} n_{i,k}$ .

## 4. Inferencia bayesiana

En cada distrito (estrato)  $i$  con  $K_i$  casillas se obtendrá una muestra de  $c_i \ll K_i$  casillas por muestreo aleatorio simple, información muestral que puede denotarse:

$$\mathcal{M}_i = \{(x_{i,1}^{(k)}, \dots, x_{i,J}^{(k)}) : k = 1, \dots, c_i\}, \quad i \in \{1, \dots, N\}, \quad (8)$$

a partir de lo cual puede obtenerse la votación total por candidato de cada distrito  $i$ , esto es  $(x_{i,1}, \dots, x_{i,J})$  donde  $x_{i,j} = \sum_{k=1}^{c_i} x_{i,j}^{(k)}$ . Con lo anterior se desea realizar inferencias respecto a los vectores de parámetros  $(\theta_{i,1}, \dots, \theta_{i,J})$ , que desde un enfoque bayesiano requiere una densidad de probabilidades a posteriori como en (1):

$$p(\theta_{i,1}, \dots, \theta_{i,J} | x_{i,1}, \dots, x_{i,J}) = \frac{1}{B} \pi(\theta_{i,1}, \dots, \theta_{i,J}) f_i(x_{i,1}, \dots, x_{i,J} | \theta_{i,1}, \dots, \theta_{i,J}). \quad (9)$$

Para una distribución de probabilidad Multinomial normalmente se utiliza una distribución a priori **Dirichlet**, véase por ejemplo Bernardo y Smith (1994), por dos razones esencialmente: primero, representa un vector de variables aleatorias cada una con soporte en el intervalo abierto  $]0, 1[$  y tales que su suma es siempre igual a 1, tal cual se requiere para los parámetros de la distribución Multinomial; segundo, es *familia conjugada* para la distribución de probabilidad Multinomial, esto es que la distribución a posteriori resulta ser también Dirichlet (aunque con parámetros actualizados de acuerdo a la información muestral):

$$\pi(\theta_{i,1}, \dots, \theta_{i,J}) = \frac{\Gamma(\sum_{j=1}^J \alpha_{i,j})}{\prod_{j=1}^J \Gamma(\alpha_{i,j})} \prod_{j=1}^J \theta_{i,j}^{\alpha_{i,j}-1} \mathbf{1}_{\{\sum_{j=1}^J \theta_{i,j} = 1\}} \quad (10)$$

donde los *hiperparámetros*  $\alpha_{i,j} > 0$  deben especificarse de acuerdo a *información a priori*. Como lo que se busca es una inferencia basada más en la información muestral que en información a priori puede optarse por utilizar  $\alpha_{i,j} = 1$  ya que en tal caso se obtiene una distribución de probabilidad *uniforme* sobre el conjunto  $\{(\theta_{i,1}, \dots, \theta_{i,J}) \in ]0, 1]^2 : \sum_{j=1}^J \theta_{i,j} = 1\}$  :

$$\pi(\theta_{i,1}, \dots, \theta_{i,J}) = (J-1)! \mathbf{1}_{\{\sum_{j=1}^J \theta_{i,j} = 1\}}, \quad (11)$$

y por lo tanto:

$$\pi(\theta_{i,1}, \dots, \theta_{i,J}) f_i(x_{i,1}, \dots, x_{i,J} | \theta_{i,1}, \dots, \theta_{i,J}) \propto \prod_{j=1}^J \theta_{i,j}^{x_{i,j}+1-1} \mathbf{1}_{\{\sum_{j=1}^J \theta_{i,j}=1\}}, \quad (12)$$

de lo cual se deduce que la distribución de probabilidad a posteriori en cada distrito  $i \in \{1, \dots, N\}$  para  $(\theta_{i,1}, \dots, \theta_{i,J})$  es **Dirichlet** $(1 + x_{i,1}, \dots, 1 + x_{i,J})$ , donde la distribución marginal a posteriori para un candidato  $j^*$  del distrito  $i$  resulta ser **Beta** $(1 + x_{i,j^*}, \sum_{j=1}^J x_{i,j} - x_{i,j^*} + J - 1)$ .

## 5. Estimación por intervalo

Con lo obtenido en la sección anterior se obtiene la distribución de probabilidad a posteriori para cada variable aleatoria  $\Theta_{i,j}$ , la proporción de votos a favor del candidato  $j$  en el distrito  $i$ , y por tanto la proporción total de votos a favor del candidato  $j$  en (4) es una combinación lineal convexa de la forma  $\Theta_j = \sum_{i=1}^N \beta_i \Theta_{i,j}$  donde  $\beta_i = n_i/n > 0$  y  $\sum_{i=1}^N \beta_i = 1$ . Definiendo  $\mu_{i,j} := \mathbb{E}(\Theta_{i,j})$  y  $\sigma_{i,j}^2 := \mathbb{V}(\Theta_{i,j})$ , por la propiedad de linealidad de la esperanza es inmediato obtener que:

$$\mu_j := \mathbb{E}(\Theta_j) = \sum_{i=1}^N \beta_i \mu_{i,j}, \quad j \in \{1, \dots, J\}, \quad (13)$$

lo cual deja claro que la esperanza de  $\Theta_j$  solo depende de las esperanzas marginales de las variables aleatorias  $\{\Theta_{i,j} : i = 1, \dots, N\}$  sin importar la posible dependencia que pudiera existir entre ellas. Sin embargo, en el cálculo de la varianza de  $\Theta_j$  sí se encuentra implicada dicha dependencia ya que aparecen covarianzas:

$$\begin{aligned} \sigma_j^2 := \mathbb{V}(\Theta_j) &= \sum_{i=1}^N \beta_i^2 \sigma_{i,j}^2 + 2 \sum_{1 \leq i < r \leq N} \beta_i \beta_r \text{Cov}(\Theta_{i,j}, \Theta_{r,j}) & (14) \\ &= \sum_{i=1}^N \sum_{r=1}^N \beta_i \beta_r \text{Cov}(\Theta_{i,j}, \Theta_{r,j}) \\ &\leq \sum_{i=1}^N \sum_{r=1}^N \beta_i \beta_r \sigma_{i,j} \sigma_{r,j} = \left( \sum_{i=1}^N \beta_i \sigma_{i,j} \right)^2 & (15) \end{aligned}$$

en donde la desigualdad (15) es consecuencia inmediata de la *Desigualdad de Cauchy-Schwarz*, ver por ejemplo Casella y Berger (2002), y proporciona el valor máximo posible de la varianza  $\sigma_j^2$ . Sea  $I_j(\alpha)$  un intervalo de probabilidad  $0 < \alpha < 1$  para  $\Theta_j$ , esto es tal que  $\mathbb{P}[\Theta_j \in I_j(\alpha)] = \alpha$ . Supongamos que para algún valor  $z_\alpha$  dicho intervalo puede expresarse en la forma:

$$I_j(\alpha) = [\mu_j - z_\alpha \sigma_j, \mu_j + z_\alpha \sigma_j] \quad (16)$$

y lo análogo para  $\Theta_{i,j}$  mediante  $I_{i,j}(\alpha) = [\mu_{i,j} - z_\alpha \sigma_{i,j}, \mu_{i,j} + z_\alpha \sigma_{i,j}]$ . Si se construye un intervalo  $I_j^*(\alpha)$  como combinación lineal convexa de los intervalos  $I_{i,j}(\alpha)$  mediante las ponderaciones  $\beta_i = n_i/n$  se

obtiene lo siguiente:

$$\begin{aligned}
 I_j^*(\alpha) &= \sum_{i=1}^N \beta_i I_{i,j}(\alpha) \\
 &= \left[ \sum_{i=1}^N \beta_i \mu_{i,j} - z_\alpha \sum_{i=1}^N \beta_i \sigma_{i,j}, \sum_{i=1}^N \beta_i \mu_{i,j} + z_\alpha \sum_{i=1}^N \beta_i \sigma_{i,j} \right] \\
 &= [\mu_j - z_\alpha \sigma_j^*, \mu_j + z_\alpha \sigma_j^*]
 \end{aligned} \tag{17}$$

donde la desviación estándar  $\sigma_j^* = \sum_{i=1}^N \beta_i \sigma_{i,j}$  es el valor máximo de  $\sigma_j$  de acuerdo a (15). En caso de que fuese válido un supuesto de independencia entre distritos entonces  $\text{Cov}(\Theta_{i,j}, \Theta_{r,j}) = 0$  y de acuerdo a (14) la desviación estándar resultante sería  $\sigma_j^\perp = \sqrt{\sum_{i=1}^N \beta_i^2 \sigma_{i,j}^2}$ , cantidad que al ser menor que  $\sigma_j^*$  da origen a un intervalo de menor longitud  $I_j^\perp(\alpha) = [\mu_j - z_\alpha \sigma_j^\perp, \mu_j + z_\alpha \sigma_j^\perp] \subset I_j^*(\alpha)$ . Lo anterior deja en claro que si el supuesto de independencia entre distritos no es correcto entonces el intervalo resultante  $I_j^\perp$  podría ser de menor longitud de la que debiera y, por lo tanto, de menor probabilidad al nivel  $\alpha$  deseado. Pero utilizar el intervalo  $I_j^*$  resultante de suponer covarianzas máximas entre distritos podría ser el otro extremo con una longitud mayor a la que debiera y por tanto con un nivel de probabilidad mucho mayor al deseado. Lo ideal sería obtener  $I_j(\alpha)$  como en (16) con el valor preciso de  $\sigma_j$  mediante (14) pero esto requería estimar  $N(N-1)/2$  covarianzas distintas con base en una cantidad de información con la que no se cuenta. Suponiendo que el valor buscado  $\sigma_j$  pertenece al intervalo  $[\sigma_j^\perp, \sigma_j^*]$  entonces debe existir un valor  $0 \leq \delta_j \leq 1$  tal que

$$\sigma_j = (1 - \delta_j) \sigma_j^\perp + \delta_j \sigma_j^* \tag{18}$$

y a partir de lo anterior el intervalo de probabilidad  $\alpha$  para  $\Theta_j$  sería:

$$\begin{aligned}
 I_j(\alpha) &= \mu_j \pm z_\alpha [(1 - \delta_j) \sigma_j^\perp + \delta_j \sigma_j^*] \\
 &= (1 - \delta_j) \mu_j + \delta_j \mu_j \pm z_\alpha (1 - \delta_j) \sigma_j^\perp \pm z_\alpha \delta_j \sigma_j^* \\
 &= (1 - \delta_j) I_j^\perp(\alpha) + \delta_j I_j^*(\alpha).
 \end{aligned} \tag{19}$$

En (19) persiste el problema de estimar  $\delta_j$  pero ya solo serían  $J$  valores (uno por candidato) además de que el valor  $0 \leq \delta_j \leq 1$  nos proporcionaría una forma de medir del grado de dependencia entre distritos por candidato (con  $\delta_j = 0$  representando independencia y  $\delta_j = 1$  covarianzas máximas), mismo que al menos podría estimarse respecto a elecciones recientes, por medio del siguiente:

### Algoritmo 1

1. Calcular los intervalos extremos  $I_j^\perp(\alpha)$  e  $I_j^*(\alpha)$  para cada  $j \in \{1, \dots, J\}$ .
2. Simular una cantidad elevada de muestras estratificadas del tamaño que se haya determinado a partir del marco muestral de una elección pasada.
3. Con base en dichas simulaciones determinar cada valor  $\delta_j$  de modo que el intervalo (19) efectivamente capture el verdadero valor  $\theta_j$  en  $100\alpha\%$  de las simulaciones del paso 2.

Se aplicó el algoritmo anterior a los resultados de las siguientes elecciones:

Candidato	Gob2011	Pre2012	Dip2015	<i>promedio</i>	<i>mediana</i>	<i>combinado</i>
Participación	0.4638	0.3722	0.5609	0.4656	0.4638	<b>0.4647</b>
NULOS	0.2622	0.0976	0.0632	0.1410	0.0976	<b>0.1193</b>
NOREG	0.3000	0.2960	0.1534	0.2498	0.2960	<b>0.2729</b>
PAN	0.3521	0.4474	0.5683	0.4559	0.4474	<b>0.4517</b>
PRI	0.2248	0.2604	0.3345	0.2732	0.2604	<b>0.2668</b>
PRD	0.2987	0.3058	0.5329	0.3791	0.3058	<b>0.3425</b>
MORENA	NA	NA	0.3272	0.3272	0.3272	<b>0.3272</b>
PT	0.2987 <sup>a</sup>	0.3058 <sup>a</sup>	0.3430	0.3158	0.3058	<b>0.3108</b>
MC	0.2987 <sup>a</sup>	0.3058 <sup>a</sup>	0.5868	0.3971	0.3058	<b>0.3515</b>
PANAL	0.2248 <sup>b</sup>	0.0852	0.3856	0.2319	0.2248	<b>0.2283</b>
PES	NA	NA	0.2056	0.2056	0.2056	<b>0.2056</b>
PH	NA	NA	0.1482	0.1482	0.1482	<b>0.1482</b>
Independientes	NA	NA	NA	NA	NA	<b>0.1769<sup>c</sup></b>

Cuadro 1: Valores de  $\delta_j$  por candidato de acuerdo al Algoritmo 1. NA = no aplica. <sup>a</sup> Valor imputado por su alianza con PRD. <sup>b</sup> Valor imputado por su alianza con PRI. <sup>c</sup> Se imputó el promedio de los valores combinados de PES y PH por ser partidos pequeños de reciente creación y que posiblemente serían el referente más cercano a lo que podría resultar de candidatos independientes que por primera vez participan en una elección de gobernador del Estado de México en la historia reciente.

- Gobernador del Estado de México en el año 2011 (Gob2011),
- Presidente de México en el año 2012, restringiéndose estrictamente a la votación en el Estado de México (Pre2012),
- Diputados federales por el Estado de México en el año 2015 (Dip2015).

En los tres casos anteriores se mapearon los resultados a la estructura de distritos locales que se utilizará en la elección de gobernador del Estado de México en el año 2017, con el objeto de que los resultados sean comparables entre sí y útiles para la elección de 2017, obteniéndose los resultados que se resumen en el Cuadro 1 con muestras estratificadas de tamaño  $c = 1,200$  proporcionales a la lista nominal por distrito y con  $\alpha = 0.95$  (intervalos de probabilidad 95%) para intervalos de proporción efectiva de votos de acuerdo a (5). La proporción de *participación* se calculó como 1 menos la proporción de abstencionismo (ABST).

En el Cuadro 1 el valor *combinado* es simplemente el punto medio entre el *promedio* y la *mediana*. Como puede apreciarse en todos los casos los valores de  $\delta_j$  son positivos lo cual significa que un supuesto de independencia entre distritos (estratos) en dichas elecciones no era realista y que en dado caso hubiese conducido a subestimaciones, en mayor o menor grado, de las varianzas  $\sigma_j^2$  y por tanto a la estimación con intervalos de probabilidad menor a la deseada de 95%, como se resume en el Cuadro 2 donde se estiman las probabilidades de cobertura de intervalos de probabilidad 95% bajo el supuesto de independencia entre distritos.

Candidato	Gob2011	Pre2012	Dip2015
Participación	39.23	46.67	33.10
NULOS	65.03	81.17	82.60
NOREG	75.10	32.07	69.63
PAN	47.07	40.27	35.57
PRI	51.37	51.93	47.27
PRD	51.27	50.70	35.33
MORENA			48.87
MC			37.03
PANAL		81.67	47.70
PES			64.13
PH			70.03

Cuadro 2: Probabilidades (en %) de que intervalos de 95% de probabilidad hubiesen capturado el verdadero resultado de la elección (cobertura) bajo el supuesto de independencia entre distritos.

## Referencias

- J.M. Bernardo y A.F.M Smith (1994) *Bayesian Theory*. Wiley (Chichester).
- G. Casella y R.L. Berger (2002) *Statistical Inference*. Duxbury (Pacific Grove).
- N.L. Johnson, S. Kotz y N. Balakrishnan (1997) *Discrete Multivariate Distributions*. Wiley (New York).
- M. Mendoza y L.E. Nieto-Barajas (2016) Quick counts in the Mexican presidential elections: A Bayesian approach. *Electoral Studies* **43**, 124–132.
- C.-E. Särndal, B. Swensson y J. Wretman (1992) *Model Assisted Survey Sampling*. Springer (New York).

## Propuesta de la Dra. Karla Beatriz Valenzuela Ocaña y el Dr. Alberto Alonso y Coria

Para estimar el porcentaje de votación que corresponde a cada partido y el porcentaje de participación ciudadana en la elección, se utilizarán los procedimientos clásicos de estimación para un muestreo aleatorio estratificado.

Como se menciona en la sección de Diseño de Muestreo, se determinó utilizar como estratos la conformación distrital local, así se tienen 45 estratos a analizar.

### I. Análisis por estrato **h**.

#### a. Estimación del número de votos por candidato

Como primer paso se estima el número de votos que obtendría cada una de las fuerzas políticas en el estrato **h**.

Sea  $v_{i,j}$  la votación por el candidato **i** en la casilla **j** y sea  $ln_j$  el listado nominal para la casilla **j**. El estimado total de participación para el candidato **i** para el estrato **h** analizado está dado por:

$$\hat{R}_i = \frac{\sum_{j=1}^n v_{i,j}}{\sum_{j=1}^n ln_j}$$

donde **n** es el número de casillas en la muestra seleccionadas del estrato **h**.

El cálculo de la varianza para el estimador  $\hat{R}_i$  se obtiene de la siguiente forma:

$$V(\hat{R}_i) \cong \frac{1-f}{nLN} \cdot \frac{\sum_{j=1}^n (v_{i,j} - \hat{R}_i ln_j)^2}{n-1}$$

donde  $f = \frac{n}{N}$ , **N** representa el número total de casillas en el estrato y  $\overline{LN}$  el listado nominal promedio de las casillas en el estrato analizado.

El intervalo de confianza para la estimación del total de participación  $\hat{R}_i$  para el candidato **i** para el estrato analizado está dado por:

$$\hat{R}_i \pm z_{\alpha/2} * \sqrt{V(\hat{R}_i)}$$

donde  $1 - \alpha$  es el nivel de confianza establecido. Para la estimación se ha definido el nivel de confianza como 95% por lo que  $z_{\alpha/2} = 1.96$

El total de votos estimado para el candidato **i** en el estrato **h** es simplemente  $\hat{R}_i LN$  donde **LN** es el listado nominal del estrato **h**.

#### b. Estimación de la participación de ciudadanos.

Es necesario estimar la participación de los ciudadanos en la elección para el estrato analizado, esto es, el número total de votos generados en el estrato.

Sea  $v_{TV,j}$  el total de votos emitidos en la casilla  $j$  incluyendo votos nulos y votos de candidatos no registrados, tenemos que  $\hat{R}_{TV}$ , el estimador de la participación ciudadana en el estrato analizado, es:

$$\hat{R}_{TV} = \frac{\sum_{j=1}^n v_{TV,j}}{\sum_{j=1}^n ln_j}$$

Este procedimiento se repite para cada uno de los  $h$  estratos.

II. Integración de la información de los estratos.

Para obtener la estimación total para cada candidato, se procede a dar el peso correspondiente a cada estrato analizado. Así pues, se agrega la información correspondiente a cada estrato  $h$ , teniendo la estimación total de votos obtenidos por el candidato  $i$ :

$$v_{T,i} = \sum_h^L \hat{R}_{i,h} LN_h$$

donde  $LN_h$  es la lista nominal del estrato  $h$  y  $\hat{R}_{i,h}$  es la participación total para el partido  $i$  en el estrato  $h$  y  $L$  es el número total de estratos.

La estimación del total de votos emitidos en el estado se calcula de manera semejante, considerando el peso correspondiente a cada estrato. Así:

$$v_{T,total\ de\ votos} = \sum_h \hat{R}_{TV,h} LN_h$$

Finalmente se tiene el intervalo de confianza para  $P_i$  como

$$P_i \pm z_{\alpha/2} * \sqrt{V_{T,i}}$$

El análisis se realiza para cada candidato  $i$ .

## Referencias

Cochran, W. G. (1977) Sampling Techniques. 3ed. Wiley&Sons, Inc.

## Propuesta del Dr. Carlos E. Rodríguez

La estimación se realizará mediante el estimador de razón combinado y técnicas de remuestreo (*Bootstrap*).

El Bootstrap es un método estadístico para estimar la distribución de muestreo de un estimador, se obtiene seleccionando  $B$  sub-muestras de la muestra original y calculando el estimador de interés con cada sub-muestra. Si se realiza un histograma con los  $B$  estimadores generados (en el caso de un estimador de una dimensión), se observaría la aproximación resultante. El número de réplicas  $B$ , debe ser suficientemente grande para lograr una buena aproximación. Una vez obtenidas las réplicas Bootstrap es posible calcular los intervalos de confianza.

Existen muchas variantes del Bootstrap, el trabajo de Efron (1979) sentó las bases, sin embargo, esta estrategia ha sido explorada por muchos autores. A continuación se describe el enfoque del Bootstrap que se implementará en el Conteo Rápido para la Elección para Gobernador en el Estado de México. En particular se usarán las ideas de Sitter (1992A) y (1992B). El primer artículo describe el método que se implementará el día de la elección, mientras que el segundo describe varias comparaciones, incluso con re-muestreo Jackknife. Estas ideas ya han sido aplicadas con éxito para realizar la estimación de la composición de la cámara de diputados en la Elección Federal de 2015, así como para la Elección Extraordinaria de Colima en 2016.

Las ideas de Sitter (1992A) aplicadas a la Elección de Gobernador en el Estado de México se describen a continuación. Sean  $N_h$  el número total de casillas instaladas en el estrato  $h$  y  $n_h$  la muestra de casillas seleccionadas (o recuperadas) en el mismo estrato. Entonces, la probabilidad de selección de una casilla en el estrato  $h$  está dada por  $f_h = \frac{n_h}{N_h}$ .

### *Esbozo del algoritmo*

Si  $N_h = n_h k_h \Leftrightarrow k_h = \frac{N_h}{n_h} = 1/f_h$  y  $m_h = f_h n_h$  (suponiendo que  $k_h$  y  $m_h$  son enteros), entonces

1. De la muestra recuperada del primer estrato (Distrito Local del Estado de México) se selecciona una sub-muestra SIN reemplazo de  $m_1$  casillas (de las  $n_1$  recuperadas).
2. Se repite el paso 1,  $k_1$  veces, obteniendo una muestra de  $n_1 = m_1 k_1$  casillas en total.
3. Se calcularán las siguientes estadísticas:
  - ✓ Estimador común del total de votos en el primer estrato para el  $i$ -ésimo candidato a gobernador

$$\hat{Y}_{1i} = N_1 \bar{y}_{1i}$$

4. Se repiten los pasos 1 a 3 para cada uno de los  $L$  estratos restantes.
5. Se calculan
  - ✓ El estimador del total de votos a favor del candidato a gobernador del partido  $i$ -ésimo en todo el estado (sumando sobre todos los estratos)

$$\hat{Y}_i = \sum_{h=1}^L \hat{Y}_{hi}$$

- ✓ Se estiman los porcentajes mediante la siguiente expresión

$$\hat{p}_i = \hat{Y}_i / \sum \hat{Y}_i$$

El paso 5 concluye una iteración Bootstrap, el objetivo es realizar  $B$  iteraciones con  $B$  suficientemente grande. Para el 4 de junio de 2017 y con cada remesa de información que se reciba se tiene pensado realizar  $B = 5,000$  iteraciones.

Es importante observar que en el paso 1, en el muestreo SIN reemplazo, la probabilidad de selección es la misma que en el diseño de muestreo original  $f_h' = \frac{m_h}{n_h} = f_h$ , esto se repite  $k_h$  veces de manera independiente para obtener una muestra de  $n_h$  casillas. Esta estrategia asegura obtener estimaciones in-sesgadas para la varianza y para el tercer momento.

En caso de que  $k_h$  o  $m_h$  no sean enteros se sigue un proceso de aleatorización para alcanzar la consistencia del algoritmo.

El algoritmo anterior es computacionalmente demandante. Sin embargo, las iteraciones del Bootstrap son independientes, así que es posible correr el algoritmo mediante procesamiento paralelo.

## Referencias

- Efron, B. (1979). "Bootstrap methods: Another look at the jackknife". The Annals of Statistics, Vol 7. No 1, pp. 1-26.
- Sitter, R. R. (1992A). "A Resampling algorithm for complex survey data". Journal of the American Statistical Association. Vol. 87. No. 419, pp. 755-765.
- Sitter, R. R. (1992B). "Comparing Three Bootstrap Methods for Survey Data". The Canadian Journal of Statistics. Vol. 20. No. 2, pp 135-154.

## IX. SIMULACROS DEL CONTEO RÁPIDO EN EL ESTADO DE MÉXICO

### a. INTRODUCCIÓN

Los simulacros del conteo rápido tienen como objetivo probar el funcionamiento de los siguientes aspectos:

- a) Los medios y sistemas para la transmisión, recepción, captura y difusión de la información electoral;
- b) El proceso operativo en campo;
- c) El ritmo de llegada de la información de las casillas en muestra;
- d) Los medios y sistemas para conocer la cobertura geográfica de la muestra;
- e) Los métodos de estimación, y
- f) La generación y envío del reporte con las estimaciones al Consejo General del IEEM.

De acuerdo a los procedimientos logísticos, los días 14 y 21 de mayo de 2017 se realizaron dos simulacros del conteo rápido. Tras evaluar los resultados de los primeros dos simulacros, la Comisión Especial para la Atención del Programa de Resultados Electorales Preliminares y los Conteos Rápidos aprobó la realización de un tercer simulacro el día 28 de mayo.

#### Tamaño de muestra y horario de cada simulacro

Simulacro	Tamaño de muestra	Fecha	Hora de inicio	Hora de corte
Primero	3352 casillas	14 de mayo	18:30 hrs	21:22 hrs
Segundo	3352 casillas	21 de mayo	15:00 hrs	18:11 hrs
Tercero	1818 casillas*	28 de mayo	10:00 hrs	12:18 hrs

\* Se utilizó una submuestra de la muestra del segundo simulacro.

Cada simulacro tuvo el siguiente esquema general de operación:

1. Para cada simulacro, COTECORA generó una muestra de casillas con votación ficticia, asociada a los partidos, alianzas y por ende a candidaturas reales para la elección del 4 de junio de 2017, claves de autenticación y un horario sugerido para reportar los datos al centro de captura.

2. En los dos primeros simulacros se diseñó una muestra que contemplara a la totalidad de los capacitadores electorales (en total 3,352 capacitadores electorales), con el fin de capacitar a cada uno de ellos. En el tercer simulacro solo se requirió la participación de 1,818 CAEs.
3. A partir de la votación ficticia de las casillas seleccionadas en muestra, se llenaron actas de escrutinio y cómputo de casillas (AECC).
4. Las AECC se hicieron llegar a la Junta Local del INE en el Estado de México para su distribución a cada Junta Distrital y a sus respectivos capacitadores electorales (CAE).
5. Cada CAE tuvo que vaciar la votación de las AECC a un formato diseñado para tal efecto (FCORA), y posteriormente vía telefónica comunicó los resultados de la votación ficticia de una casilla de su ARE para su captura en el Sistema de Captura del Conteo Rápido (SICCORA).
6. Los capturistas atendieron los procedimientos definidos para el reporte y captura de datos, según se indica en el apartado "b) Durante la Jornada Electoral" de los procedimientos logísticos.
7. La Junta Local del INE pudo dar seguimiento al avance de las casillas reportadas por cada distrito, con el fin de que pudieran conocer cualquier contratiempo y, en su caso, tomar las acciones necesarias para garantizar el flujo de la información.
8. El COTECORA recibió remesas o cortes de la información recolectada por el SICCORA cada 10 o 15 minutos.
9. El COTECORA monitoreó la cobertura de la muestra en cada remesa, el ritmo de llegada de la información, así como las tendencias preliminares estimadas conforme a los criterios científicos.
10. Una vez que las estimaciones se estabilizaron y/o se contaba con una cobertura suficiente de la muestra, el COTECORA realizó una estimación final y la asentó en un informe de resultados --siguiendo los criterios científicos--, misma que se presentó al final del simulacro ante la Comisión Especial para la Atención del Programa de Resultados Electorales Preliminares y los Conteos Rápidos.
11. Finalmente, el COTECORA realizó un reporte de observaciones de cada simulacro.
12. El domingo 14 y el viernes 26 de mayo, se hicieron pruebas del protocolo de generación de muestra.

## **b. SOBRE LA COBERTURA DE LA MUESTRA**

Al corte del primer simulacro se consiguió una cobertura de 81.77% de las 3,352

casillas en muestra. Para el segundo simulacro, este nivel aumentó hasta 96.18%. En el tercer simulacro, el porcentaje de casillas reportadas disminuyó a 94.44%, pero vale la pena destacar que este tuvo una menor duración que los primeros dos simulacros. En general, no se manifestó un problema con el número de casillas capturadas.

En particular, es relevante hacer notar que los distritos locales 9 y 37 reportaron una baja cobertura. En el primero, segundo y tercer simulacro, el distrito local 9 reportó 67.62, 80 y 70.18% de las casillas en muestra, respectivamente. Por su parte, el distrito local 37 reportó 71.21, 86.36 y 61.11%, respectivamente. Los distritos federales 16 y 36 también reportaron una baja cobertura durante los tres simulacros. En el primero, segundo y tercer simulacro, el distrito federal 16 reportó 75%, 85.71% y 64.44% de las casillas en muestra, respectivamente. El distrito federal 36 reportó 62%, 79% y 62.5%, respectivamente.

El distrito federal 36 (con cabecera en Tejupilco) se traslapa en su mayoría con el distrito local 9, mientras que el distrito federal 16 (con cabecera en Ecatepec) se traslapa con el distrito local 37 (Tlalnepantla).

### **c. SOBRE EL CENTRO DE CAPTURA**

En el primer simulacro se contó con 84 capturistas activos. El promedio de casillas por capturista fue de 36 llamadas. En el segundo simulacro se contó con 98 capturistas activos. El promedio de casillas por capturista fue de 33.8 llamadas, con un rango intercuartiles de 27 a 39 casillas. La duración promedio de las llamadas fue de 2:36, con un rango intercuartiles de 2:06 a 2:55 minutos. En el tercer simulacro se contó con 85 capturistas activos. El promedio de casillas por capturista fue de 20.2 llamadas, con un rango intercuartiles de 19 a 22 casillas capturadas. La duración promedio de las llamadas fue de 2 minutos 39 segundos, con un rango intercuartiles de 2:08 a 2:56.

En resumen, los tres simulacros evidenciaron que el centro de captura del conteo rápido era capaz de procesar más de mil casillas en una hora, y hasta 1,585 en hora y media.

### **d. SOBRE LAS ESTIMACIONES DE LOS SIMULACROS**

Al final de cada simulacro, el COTECORA realizó una estimación final del ejercicio, calculando un intervalo de confianza para cada candidatura, candidatos no registrados, votos nulos y tasa de participación electoral. De acuerdo a los criterios científicos, se realizaron tres estimaciones con metodologías estadísticas distintas, pero utilizando la misma remesa de datos, y al final se combinaron los resultados para producir una sola estimación.

Estas estimaciones se asentaron en un informe final y se presentaron ante la Comisión Especial para la Atención del Programa de Resultados Electorales

Preliminares y los Conteos Rápidos en un formato similar al que se entregará el día de la jornada.

El primer simulacro tuvo una votación ficticia que producía un resultado muy cerrado entre tres candidaturas. El segundo y tercer simulacros, por otro lado, utilizaron una votación ficticia en la que una candidatura tenía una ventaja clara sobre el resto.

En la mayoría de los casos, el intervalo de confianza estimado “contuvo” la proporción de votos simulada. Sin embargo, en cada simulacro hubo intervalos que no contenían las proporciones simuladas ya sea por algunas décimas o centésimas de punto porcentual. Esto ocurrió en cuatro casos en el primer simulacro, en tres en el segundo simulacro, y en dos casos en el tercero. Cabe destacar que con cada simulacro aumentó el número de intervalos estimados acertados. Por otro lado, la candidatura E –misma que correspondió al mismo partido político en los tres casos-- resultó subestimada en los tres simulacros.

#### **e. SOBRE CALIDAD DE LA INFORMACIÓN RECIBIDA EN LOS SIMULACROS**

Al finalizar cada simulacro, se compararon los datos de cada casilla de la muestra recibida en el SICCORA contra la muestra de votos simulada previamente por COTECORA. Cuando la casilla fue reportada y capturada sin error, la diferencia entre una y otra base es cero. Si el dato recibido es mayor (menor) al dato simulado, la diferencia es positiva (negativa).

La mayoría de los rubros de votación tuvieron diferencias promedio positivas en los tres simulacros. Por un lado, la candidatura B, en los tres simulacros, presentó diferencias positivas y significativas, de entre 1 y 2.9 votos por casilla, mientras que la candidatura E presentó diferencias negativas de entre -0.5 y -0.67 votos por casilla. En razón de que las candidaturas B y E correspondieron a la misma candidatura en los tres simulacros, se alertó por un posible sesgo en la información recibida.

#### **f. SOBRE LAS FUENTES DE ERROR SUGERIDAS POR LOS SIMULACROS**

La confiabilidad y precisión de las estimaciones del conteo rápido depende de manera muy importante de la cantidad y calidad de los datos recibidos. Así pues, los simulacros permitieron detectar y corregir o al menos mitigar posibles fuentes de error en la captura y transmisión de la información. Entre los problemas detectados se encuentran:

- a) No reporte o cobertura desbalanceada. Si algunos distritos o tipos de casilla reportan una proporción relativamente baja de datos.

- b) Error en el acta de escrutinio y cómputo. Si el acta tiene algún dato ilegible o no disponible, mismo que se subsana en el cómputo distrital o en el recuento.
- c) Error en la casilla a reportar o la clave de autenticación. Si el CAE no conoce con antelación la casilla o casillas que deberá reportar, así como la clave de autenticación, no podrá transmitir el resultado.
- d) Error en el reporte. Si el CAE reporta al centro de captura datos distintos a los asentados en el acta de escrutinio y cómputo.
- e) Error en la captura. Si el capturista transcribe datos distintos a los que le dicta el CAE.
- f) El SICCORA no almacene la información de forma adecuada, e introduzca desviaciones con respecto a la misma.

Para minimizar estas fuentes de error se recomendaron las siguientes medidas:

- a) Distribuir con antelación a la jornada electoral, la información encriptada de la muestra y claves de autenticación.
- b) Que el SICCORA valide o verifique cualquier casilla que reporte un total de votos mayor a la lista nominal (más 20, para dar cabida a los representantes de partido).
- c) Doble captura ciega de la identificación de casilla, así como de la votación reportada.
- d) Reforzar y mejorar el protocolo empleado por los capturistas y CAEs para la captura de los resultados de las casillas.
- e) Solicitar al INE que los CAE graben las llamadas al centro de captura y resguarden el formato de recopilación de datos utilizados.
- f) Establecer protocolos de seguridad e integridad en el flujo de información de la votación de las casillas en muestra.

## **X. SELECCIÓN DE LA MUESTRA.**

### **1. Consideraciones generales para la selección de la muestra**

La generación de la muestra de casillas para el conteo rápido se realizó en un acto público el viernes 2 de junio de 2017 a las 10:00 horas. En el acto estuvo presente un notario público que dio fe del desarrollo del protocolo desde la instalación del software requerido hasta la obtención y resguardo de la muestra definitiva.

El acto público, contó con la presencia de los consejeros electorales, representantes de los partidos políticos y de medios de comunicación. Consistió

en una breve presentación de los conteos rápidos, de la preparación en el equipo de cómputo del entorno para la generación de la muestra, de la propia generación de la muestra, del borrado seguro de archivos, y por último, de la entrega resguardo y transmisión al INE del contenido de la misma.

En la etapa de preparación del entorno hubo una problemática con el equipo de cómputo utilizado y hubo que recurrir a un equipo de un miembro del COTECORA.

Una vez habilitado el equipo de cómputo se procedió a generar la semilla a utilizar para la utilización del programa pseudoaleatorio de generación.

- 1.- Se formaron tres parejas de consejeros del IEEM
- 2.- Cada pareja, utilizando una tómbola, generó un número de 6 dígitos (el primero diferente de cero)
- 3.- Cada pareja ingresó el número generado en la computadora, dos veces. Cada pareja ingresó su número por separado.

Una vez realizado lo anterior se generó la muestra la cual arrojó 40 archivos correspondientes a las casillas seleccionadas de cada uno de los distritos federales.

Se procedió a cifrar la información y transmitir en ese momento a las 40 juntas distritales del INE para que procedieran a informar a los CAES que tendrían la responsabilidad de transmitir el día de la jornada electoral la información de los resultados de las casillas.

En el acto se generaron dos discos compactos no regrabables conteniendo la muestra. Uno que sería utilizado por la Unidad de Informática para ser cargado el día de la jornada electoral en el sistema de cómputo, puesto que se requería de esta información para validar que los resultados que se transmitieran correspondieran a una de las casillas de la muestra.

El segundo disco fue uno con funciones de testigo, y se resguardó en presencia de notario. Este disco se acompañó con un tercero, pero éste último contenía los insumos que se utilizaron para la generación de la muestra. Asimismo, se resguardó el equipo de cómputo utilizado, garantizando así que no hubiera posibilidades de extraer el contenido de la muestra indebidamente.

## **XI. DIA DE LA JORNADA ELECTORAL.**

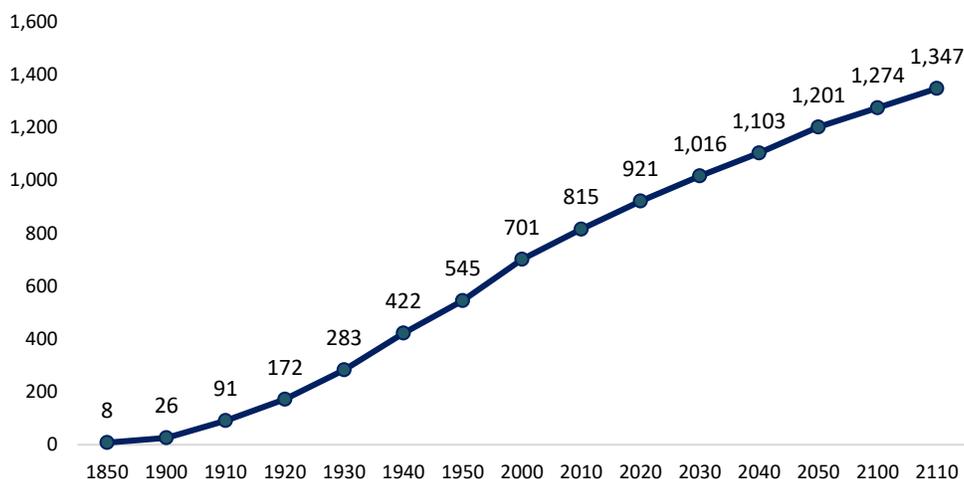
### **1. Sobre la recepción de información**

El centro de captura del IEEM contó con 86 capturistas para recibir los reportes de los CAE desde las 18:00 horas del 4 de junio.

De acuerdo a lo programado, la información de las casillas reportadas fluyó al COTECORA en forma de remesas o cortes. Esto se hizo con la finalidad de que en todo momento los integrantes de COTECORA tuvieran los mismos datos para

poder hacer estimaciones comparables, producto de los tres procedimientos de estimación. Las remesas se generaron cada diez minutos. La primera remesa con información de casillas capturadas llegó a las 18:50 horas, en donde se recibieron 8 casillas. La serie de tiempo con el flujo de información del número de casillas capturadas, se muestra a continuación.

**Número de casillas recabadas  
(corte a las 21:10 horas del 4 de junio de 2017)**



Conforme a lo señalado en los criterios científicos del conteo rápido, se esperaba que con la información de una **muestra objetiva** de 1,200 casillas se podrían obtener estimaciones con una precisión de máximo 0.5% y con un nivel de confianza del 95%. En el mismo documento, se estableció que se generaría una **muestra ampliada** de 1,818 casillas para poder soportar una tasa de no respuesta o reporte tardío de 34%:  $1,200 / 0.66 = 1,818$  casillas.

A las 21:10 horas del 4 de junio de 2017 se contaba con 1,374 casillas distribuidas en los 45 estratos del diseño muestral. Esta muestra tenía un número de casillas suficiente que permitía realizar las estimaciones con el nivel de precisión y confianza establecidos previamente, además de que las estimaciones hasta ese momento se apreciaban razonablemente estables. Así pues, con la información correspondiente al corte de las 21:10, el COTECORA realizó su estimación final y procedió a realizar el informe de resultados correspondiente.

## 2. Sobre la cobertura de la muestra

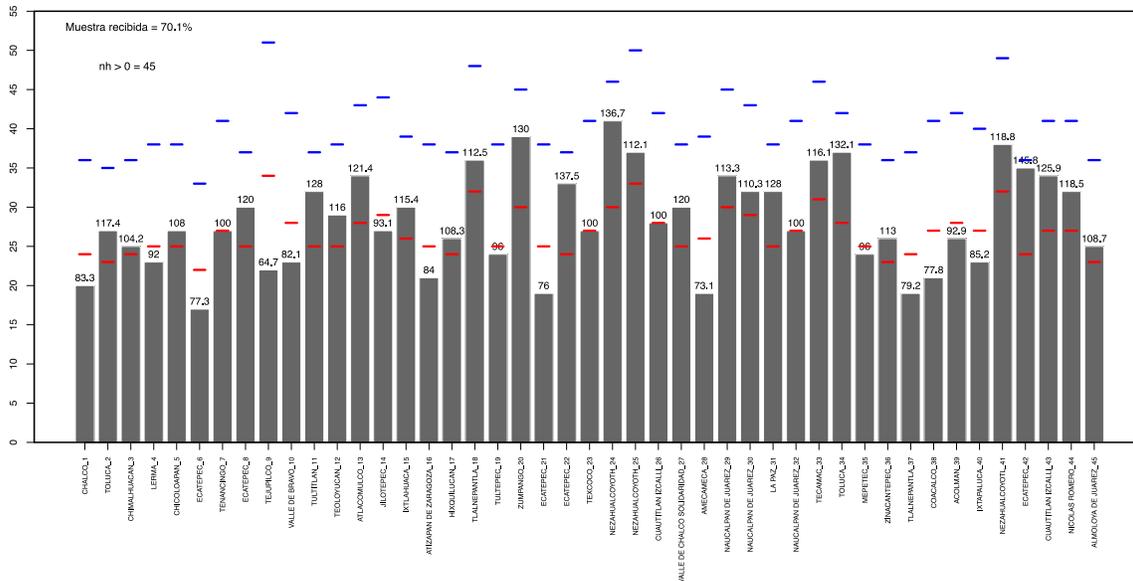
Como se estableció en los criterios científicos, las estimaciones se realizaron utilizando como estratos a los 45 distritos locales de la entidad. La cobertura de la muestra por distrito local hasta las 21:10 horas, se puede observar en la siguiente tabla:

Casillas recibidas por distrito local  
(corte a las 21:10 hrs del 4 de junio de 2017)

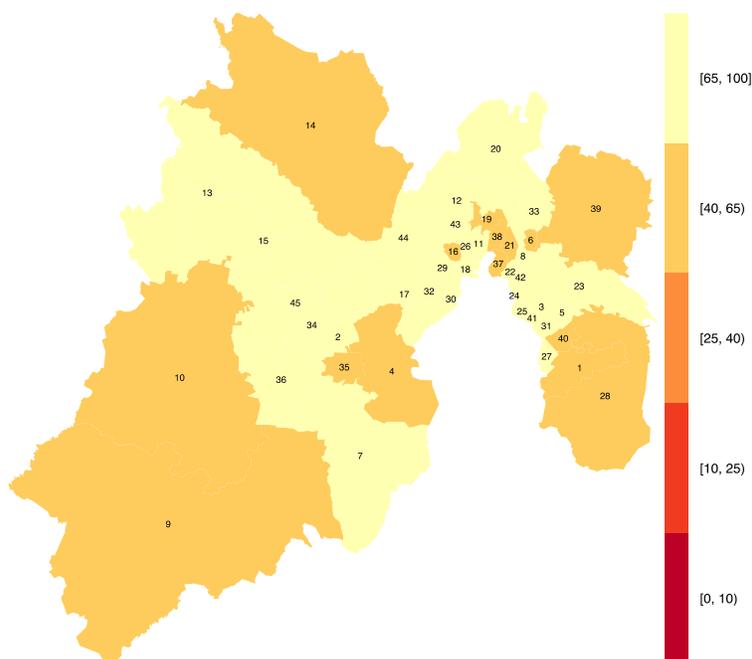
Distrito Local	Casillas reportadas	Casillas en muestra	% avance
1	23	36	63.89
2	27	35	77.14
3	26	36	72.22
4	25	38	65.79
5	29	38	76.32
6	17	33	51.52
7	29	41	70.73
8	31	37	83.78
9	25	51	49.02
10	26	42	61.90
11	33	37	89.19
12	30	38	78.95
13	34	43	79.07
14	27	44	61.36
15	31	39	79.49
16	23	38	60.53
17	28	37	75.68
18	38	48	79.17
19	24	38	63.16
20	42	45	93.33
21	24	38	63.16
22	33	37	89.19
23	29	41	70.73
24	43	46	93.48
25	38	50	76.00
26	29	42	69.05
27	33	39	84.62
28	19	39	48.72
29	37	45	82.22
30	32	43	74.42
31	33	38	86.84
32	30	41	73.17
33	40	46	86.96
34	37	42	88.10
35	26	38	68.42
36	28	36	77.78
37	21	37	56.76

38	23	41	56.10
39	31	42	73.81
40	27	40	67.50
41	39	49	79.59
42	35	36	97.22
43	34	41	82.93
44	32	41	78.05
45	26	36	72.22
<b>TOTALES</b>	<b>1347</b>	<b>1818</b>	<b>74.09</b>

La siguiente gráfica ilustra la cobertura y balance de la muestra recibida por distrito local en el corte de las 21:00 horas del 4 de junio. Las marcas rojas señalan el porcentaje de muestra que se debió haber alcanzado, tomando como referencia la muestra objetivo de 1,200 casillas, y las marcas azules toman como referencia la muestra ampliada de 1,818 casillas.



El siguiente mapa ilustra el avance de la muestra en cada distrito local al 21:00 horas. En ese corte ya se contaba 1274 casillas, lo cual ya cubría la muestra objetivo de 1,200 casillas y representaba el 70% de la muestra ampliada de 1818 casillas.



En cuanto a la ubicación de casillas rurales y urbanas, la muestra final tuvo la siguiente composición: 23.5% de casillas rurales y 76.5% de casillas urbanas.

UBICACION  CASILLA	Frecuencia	Porcentaje
URBANA	1,030	76.47
RURAL	317	23.53
Total	1,347	100.00

Mientras que el marco muestral integrado por el total de casillas a instalarse tenía esta distribución:

UBICACION  CASILLA	Frecuencia	Porcentaje
URBANA	13,965	75.06
RURAL	4,641	24.94
Total	18,606	100.00

### 3. Sobre las estimaciones

Para dar mayor certeza a las estimaciones, los integrantes del COTECORA utilizaron tres métodos estadísticos distintos para estimar los resultados de la elección:

- Método Bayesiano, a cargo del Dr. Arturo Erdely Ruiz.
- Método Clásico vía las fórmulas para estimadores de razón, derivadas mediante las técnicas de muestreo probabilístico, a cargo de la Dra. Karla Beatriz Valenzuela Ocaña y el Dr. Alberto Alonso y Coria.
- Método Clásico vía métodos de remuestreo, a cargo del Dr. Carlos Erwin Rodríguez.

Los resultados obtenidos por las tres estimaciones se integraron promediando la media y la mediana de los intervalos obtenidos con los tres métodos, de acuerdo a la metodología previamente establecida en los criterios científicos. La tabla siguiente ilustra las tres estimaciones finales, así como los intervalos combinados:

**CONTEO RÁPIDO ESTADO DE MÉXICO 2017  
COMBINACION DE INTERVALOS ESTIMADOS**

CANDIDATURA	COMBINACION			ERWIN RODRIGUEZ			ARTURO ERDELY			ALBERTO ALONSO / KARLA VALENZUELA		
	LI	LS	Precisión	LI	LS	Precisión	LI	LS	Precisión	LI	LS	Precisión
JOSEFINA VÁZQUEZ MOTA	10.99	11.57	0.29	11.00	11.55	0.28	10.97	11.63	0.33	11.00	11.56	0.28
ALFREDO DEL MAZO MAZA	32.75	33.59	0.42	32.75	33.60	0.43	32.77	33.52	0.38	32.71	33.61	0.45
JUAN MANUEL ZEPEDA	17.60	18.28	0.34	17.60	18.27	0.33	17.61	18.29	0.34	17.60	18.29	0.34
OSCAR GONZÁLEZ YÁÑEZ	1.03	1.13	0.05	1.04	1.12	0.04	1.01	1.17	0.08	1.04	1.12	0.04
DELFINA GÓMEZ ÁLVAREZ	30.73	31.53	0.40	30.77	31.52	0.38	30.69	31.54	0.42	30.73	31.54	0.41
MARÍA TERESA CASTELL	2.15	2.27	0.06	2.16	2.26	0.05	2.13	2.30	0.09	2.15	2.26	0.05
Candidatos No Registrados	0.08	0.11	0.01	0.08	0.10	0.01	0.08	0.13	0.02	0.08	0.10	0.01
Votos Nulos	2.95	3.29	0.17	2.94	3.32	0.19	3.02	3.20	0.09	2.91	3.30	0.20
Participación	53.31	54.15	0.42	52.89	53.95	0.53	53.38	54.17	0.39	53.42	54.26	0.42

Como puede apreciarse en la tabla anterior, los tres métodos de estimación resultaron consistentes entre sí, lo cual contribuyó a que los integrantes de COTECORA acordaran realizar una estimación final. Estas fueron las estimaciones reportadas en el informe de resultados que se entregó --ante notario público-- al Consejero presidente del Consejo General del IEEM, con corte las 21:10 horas y con 1,347 casillas recibidas:

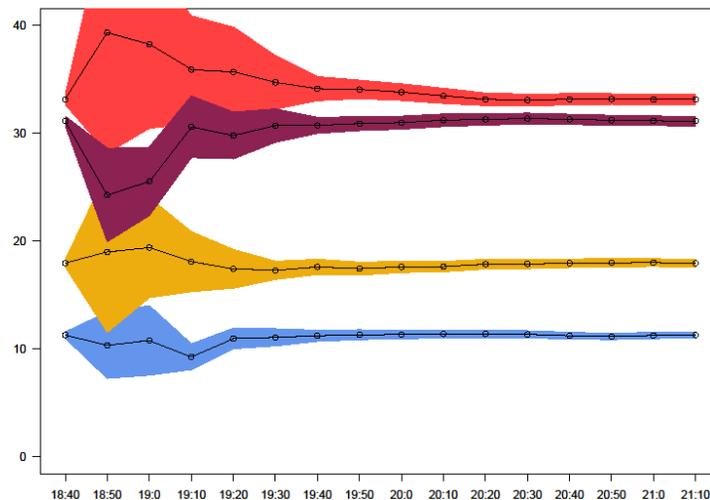
#### Estimación de tendencias de la votación

Candidatura	Límite inferior	Límite superior	Precisión
JOSEFINA EUGENIA VÁZQUEZ MOTA	10.99	11.57	0.29
ALFREDO DEL MAZO MAZA	32.75	33.59	0.42
JUAN MANUEL ZEPEDA HERNÁNDEZ	17.60	18.28	0.34
OSCAR GONZÁLEZ YÁÑEZ	1.03	1.13	0.05
DELFINA GÓMEZ ÁLVAREZ	30.73	31.53	0.40
MARÍA TERESA CASTELL DE ORO PALACIOS	2.15	2.27	0.06
Candidatos No Registrados	0.08	0.11	0.015
Votos Nulos	2.95	3.29	0.17
Participación	53.31	54.15	0.42

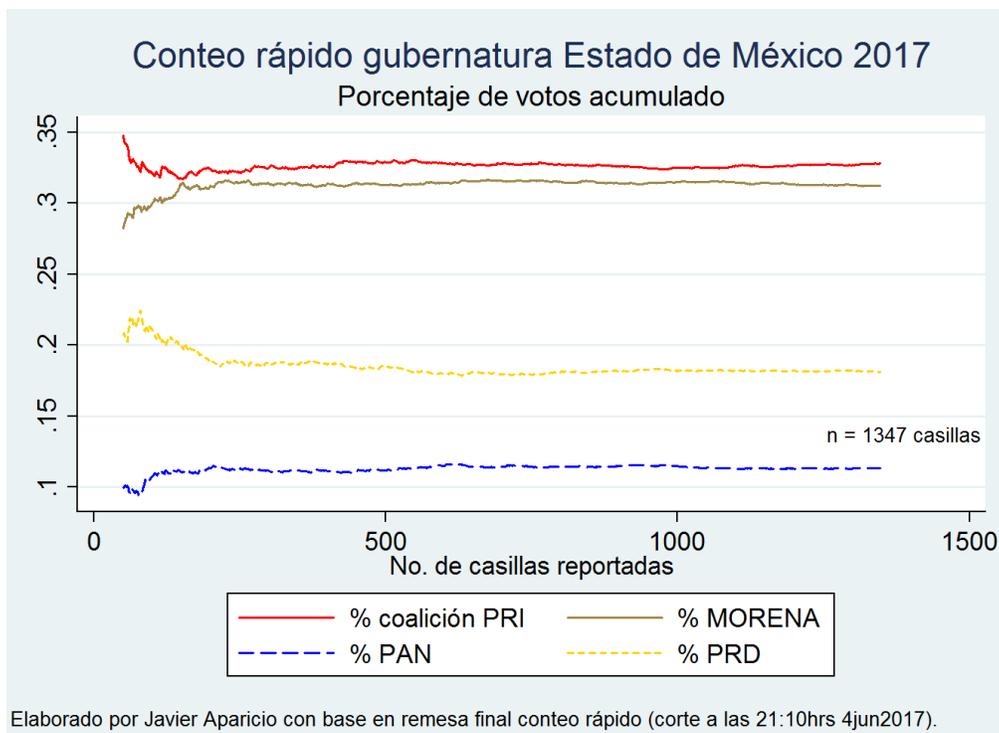
*Nota: estimación con 1,347 casillas recibidas a las 21:10 horas del 4 de junio de 2017.*

La gráfica siguiente muestra los intervalos del porcentaje de votación estimados para los primeros cuatro lugares de la contienda durante la noche de la jornada electoral, con el siguiente código de colores: Alfredo del Mazo Maza (rojo), Delfina Gómez Álvarez (morado), Juan Manuel Zepeda Hernández (amarillo) y Josefina Eugenia Vázquez Mota (azul).

**Intervalos estimados para candidaturas selectas  
cortes 18:40 y 21:10 horas del 4 de junio de 2017**



La gráfica siguiente muestra las estimaciones puntuales del porcentaje de votación acumulada para los primeros cuatro lugares de la contienda conforme avanzaba el número de casillas recibidas la noche de la jornada electoral. Como puede apreciarse en esta y la gráfica anterior, las estimaciones se estabilizaron relativamente a partir de contar con al menos 1000 casillas capturadas, esto es, alrededor de las 20:30 horas.



#### 4. Sobre la supervisión logística y coordinación con el INE

Previo a la jornada electoral, en los tres simulacros realizados, y durante la jornada misma, el Dr. Javier Aparicio se encargó de la supervisión logística del conteo rápido. Para tal efecto, mantuvo en contacto con las autoridades correspondientes en el INE para garantizar la recepción de la información sobre la muestra y dar seguimiento al avance en la información recibida. Adicionalmente, revisó en todo momento que la cobertura obtenida fuera la adecuada para poder generar las estimaciones presentadas anteriormente.

## XII. ANÁLISIS POST-JORNADA ELECTORAL

Posteriormente a la jornada electoral se realizó la comparación de las estimaciones con el resultado final arrojado tanto por el PREP como por los cómputos distritales.

Instrumentos del IEEM para estimar o calcular el porcentaje de votos en favor de cada candidato a Gobernador (Comparativo)

INSTRUMENTO DEL IEEM	FECHA	HORA	NÚMERO DE ACTAS USADAS
<b>CONTEO RÁPIDO</b>	4 DE JUNIO	9:50 p.m.	1,347
<b>PREP</b>	5 DE JUNIO	1 p.m.	18,173
<b>CÓMPUTOS DISTRITALES</b>	9 DE JUNIO	10 a.m.	18,606

### Elección para gubernatura Estado de México 2017 Comparativo del porcentaje de votos para cada candidatura

Candidaturas	Límite inferior	Límite superior	PREP	Cómputos Distritales
JOSEFINA EUGENIA VÁZQUEZ MOTA	10.99	11.57	11.29	11.28
<b>ALFREDO DEL MAZO MAZA</b>	<b>32.75</b>	<b>33.59</b>	<b>33.72</b>	<b>33.69</b>
JUAN MANUEL ZEPEDA HERNÁNDEZ	17.6	18.28	17.79	17.89
OSCAR GONZÁLEZ YÁÑEZ	1.03	1.13	1.08	1.08
<b>DELFINA GÓMEZ ÁLVAREZ</b>	<b>30.73</b>	<b>31.53</b>	<b>30.82</b>	<b>30.91</b>
MARÍA TERESA CASTELL DE ORO PALACIOS	2.15	2.27	2.13	2.14
CANDIDATOS NO REGISTRADOS	0.08	0.11	0.13	0.11
VOTOS NULOS	2.95	3.29	3.03	2.89
PARTICIPACIÓN	53.31	54.15	52.5	53.71

La oportunidad con la que se presentaron las estimaciones del Conteo Rápido fue tal como se había planeado y sirvió para abonar a la claridad y transparencia del proceso electoral, en un momento en el que no se tenía información confiable.

Como puede observarse, los intervalos del Conteo Rápido cubren prácticamente todos los resultados dados por los cómputos distritales, a excepción de la alianza en torno a la candidatura de Alfredo del Mazo Maza, que no se cubre por 10 décimas de punto porcentual.

El Conteo Rápido y el PREP, no son completamente comparables ya que:

- El PREP abarcó alrededor del 97.7 % de las actas totales, dejando un 2.3 % fuera (esto es algo común). En cambio, el Conteo Rápido busca estimar los resultados para el 100% de las casillas.
- Observando sólo las celdas de la base de datos que tienen datos acerca de la votación, se tienen 409,332 celdas. De las cuales 46,809, aproximadamente 11.4%, tienen datos faltantes (clasificados como SIN DATO o ILEGIBLE). Por otro lado, la base de datos del Conteo Rápido no tiene datos faltantes.
- Poco menos de 90% de las casillas con datos faltantes eran de alguna de las combinaciones de votación a favor de la alianza en torno a Alfredo del Mazo Maza.
- El PREP siempre subestima la participación ya que compara la votación total de 97.7 % de casillas (en este caso), contra la lista nominal del 100% de las casillas. El Conteo Rápido estima el número total de votos para todas las casillas instaladas y lo compara contra la lista nominal del 100% de las casillas,

Estos detalles, hacen que estos dos instrumentos no sean del todo comparables.

Por último, no es omiso señalar, que una comparación con lo arrojado por el PREP de la base de datos transmitida señaló que los errores de captura fueron de magnitud menores y con una distribución neutral.

### **XIII. CONCLUSIONES Y RECOMENDACIONES**

El ejercicio del Conteo Rápido llevado a cabo el día 4 de junio de 2017 resultó exitoso ya que permitió dar a conocer al público de manera oportuna y con precisión estadística las tendencias de la votación para cada candidatura. A partir de la muestra recuperada y de acuerdo a las estimaciones realizadas por el Comité Técnico Asesor del Conteo Rápido, se encontró una diferencia estadísticamente significativa entre el primer y segundo lugar de la elección para gobernador del Estado de México.

#### **RECOMENDACIONES**

- Se deben realizar al menos dos simulacros de selección de muestra completos, por lo menos uno bajo las mismas condiciones que se tendrán el día de la selección de la muestra definitiva. Además, se deberá usar una computadora o laptop para generar la muestra y desde esa misma laptop, enviar los correos a las juntas distritales. Una computadora o laptop adicional, debe ponerse en reserva en caso de que la primera tenga algún fallo.
- Que tanto en los simulacros como el día de la jornada electoral se graben las llamadas que realicen los CAEs al centro de captura para el conteo rápido para que la fidelidad de la información transmitida y capturada sea verificable, de resultar necesario.

- Para futuros conteos rápidos incluir dentro de los criterios científicos y metodológicos para tal fin pruebas estadísticas formales para: a) determinar si existe una diferencia estadísticamente significativa entre los dos candidatos punteros, y b) determinar si las tendencias de votación pueden considerarse estables con la muestra recibida al momento de hacer la estimación final.