



PRIMER INFORME QUE PRESENTA EL CONSEJERO PRESIDENTE A LOS INTEGRANTES DEL CONSEJO GENERAL RESPECTO A LAS ACTIVIDADES QUE HA DESARROLLADO EL COMITÉ TÉCNICO ASESOR PARA EL CONTEO RÁPIDO (COTECORA) 2012

ABRIL 2012

ANTECEDENTES

- I. El Comité Técnico Asesor para el Conteo Rápido (COTECORA) fue creado mediante el Acuerdo del Consejo General número CG149/2012, aprobado en la sesión ordinaria del 14 de marzo de 2012. Instituyéndose su funcionamiento a partir del 15 de marzo y hasta el 31 de agosto de 2012.
- II. De conformidad con lo dispuesto en el punto Cuarto del Acuerdo del Consejo General referido, el Comité tiene entre sus funciones rendir informes mensuales al Consejo General, motivo por el cual, se presenta el informe relativo a las actividades realizadas durante el periodo que va del 21 de marzo al 16 de abril.
- III. Tal como se detalla en el punto Sexto del Acuerdo del Consejo General citado, el Comité se encuentra conformado por los siguientes miembros:
 1. Dra. Guillermina Eslava Gómez, Profesor titular "A", tiempo completo en la Facultad de Ciencias de la Universidad Nacional Autónoma de México.
 2. Mtra. Patricia Isabel Romero Mares, Técnico Académico Titular "B", tiempo completo definitivo en el Departamento de Probabilidad y Estadística del Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas de la Universidad Nacional Autónoma de México.
 3. Dr. Rubén Hernández Cid, Profesor Titular del Departamento de Estadística del Instituto Tecnológico Autónomo de México.
 4. Dr. Manuel Mendoza Ramírez, Director del Centro de Estadística Aplicada del Instituto Tecnológico Autónomo de México.
 5. Dr. Raúl Rueda Díaz del Campo, Investigador del Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas de la Universidad Nacional Autónoma de México.
 6. Ing. René Miranda Jaimes, Secretario Técnico del Comité Técnico Asesor para el Conteo Rápido.

INFORME

I. Sesión de Instalación

El día 21 de marzo de 2012 en las salas 1 y 2 de Consejeros ubicadas en la planta baja del edificio "A" del Instituto Federal Electoral, se realizó la Sesión de Instalación del Comité Técnico Asesor para el Conteo Rápido. En dicha sesión se contó con la presencia del Consejero Presidente, el Dr. Leonardo Valdés Zurita; el Secretario Ejecutivo, Lic. Edmundo Jacobo Molina; los Consejeros Electorales, Dr. Lorenzo Córdova Vianello y Dr. Sergio García Ramírez; el Lic. Rodrigo Moreno González, asesor del Consejero del Poder Legislativo del PAN; el Lic. Elliot Báez Ramón, asesor del Consejero del Poder Legislativo del PRI; el Lic. Alejandro Álvarez Fernández, asesor del Consejero del Poder Legislativo del PRD; la Lic. Lilyana Barajas R., asesora de la representación de Movimiento Ciudadano; el Lic. Everardo Rojas Soriano, representante suplente del PAN; el Lic. Juan Antonio Mora

García, por ausencia del representante del PRI; el Lic. Jaime Miguel Castañeda Salazar, asesor del representante del PRD; el Lic. Ricardo Cantú Garza, representante del PT; la Dra. Leticia Amezcua, asesora de la representación del PVEM; el Lic. Luis Antonio González Roldán, representante del PANAL; así como todos los miembros y el Secretario Técnico del Comité.

En la sesión tomaron la palabra el Consejero Presidente, Dr. Leonardo Valdés Zurita; el Secretario Técnico del Comité, Ing. René Miranda Jaimes; y, en representación de este órgano, el Dr. Rubén Hernández Cid. Los tres coincidieron en la trascendencia que este tipo de ejercicios ha tenido en procesos electorales pasados. Además de mostrarse plenamente convencidos de que el conteo rápido es un instrumento que contribuye a la transparencia y la certeza del proceso electoral, al ofrecer a la ciudadanía, la misma noche de la elección, una estimación preliminar de la votación de Presidente de los Estados Unidos Mexicanos.

II. Reuniones de trabajo

El Comité Técnico Asesor para el Conteo Rápido se ha reunido en sesiones ordinarias en tres ocasiones: el 30 de marzo y el 9 y 16 de abril de 2012.

En la primera sesión ordinaria se revisó y aprobó el Plan de Trabajo del Comité; se analizaron los ejercicios de Conteo Rápido realizados en 2000 y 2006; se expuso la evolución del Padrón Electoral, el Sistema de Información de la Jornada Electoral y el sistema de captura del Conteo Rápido.

En la segunda sesión ordinaria se analizó la solicitud del Partido Verde Ecologista de México para ampliar el alcance del ejercicio de conteo rápido a las elecciones de Senadores y Diputados por el principio de mayoría relativa; se revisó el Programa de Operación Logística del Conteo Rápido y el Comité remitió sus comentarios y observaciones sobre el mismo y el proceso de Reseccionamiento, con el apoyo de la Dirección Ejecutiva de Organización Electoral (DEOE) y la Dirección Ejecutiva del Registro Federal de Electores (DERFE), respectivamente.

En la tercera sesión ordinaria se hizo el planteamiento de los temas que se revisarán en la sesión de trabajo a realizarse con las representaciones de los partidos políticos y los Consejeros del Poder Legislativo, se revisaron los avances que se han hecho en la construcción del diseño muestral, se plantearon los filtros que se aplicarán en las remesas y se hicieron observaciones al esquema logístico del Conteo Rápido.

III. Plan de trabajo del Comité Técnico Asesor para el Conteo

Se elaboró y aprobó el Plan de Trabajo que regirá las actividades a desarrollar durante el periodo en que se llevarán a cabo los trabajos que el Consejo General encomendó a este grupo de especialistas. En el mismo, se establecen los objetivos, el marco normativo y la naturaleza jurídica, los integrantes, el organigrama, las funciones, el formato y el cronograma de actividades.

En dicho Plan se especifica que el objetivo del Comité es ofrecer en tiempo y forma, con validez estadística, un pronóstico sobre los resultados electorales de la elección de Presidente de los Estados Unidos Mexicanos.

Para cumplir con sus funciones, el Comité contará con el acompañamiento de la Presidencia del Consejo General, la Secretaría Ejecutiva y la Unidad de Servicios de Informática (UNICOM). Además de estar facultado para coordinarse con diversas áreas ejecutivas del Instituto Federal Electoral.

Con el propósito de alcanzar el objetivo que se estableció, se delimitaron las funciones en las actividades que se mencionan a continuación:

1. Establecer los criterios científicos del diseño muestral.
2. Acordar conjuntamente con el Instituto la logística del levantamiento de la muestra.
3. Establecer la metodología para la estimación de resultados.
4. Determinar los niveles de confianza de los estimadores.
5. Entregar el informe de los resultados el día de la Jornada Electoral en los términos que se acuerden con el Consejero Presidente.

IV. Avances en el desarrollo de las actividades del Comité Técnico Asesor para el Conteo Rápido

1. **Análisis de los ejercicios previos de Conteo Rápido en los procesos electorales federales de 2000 y 2006.** Se resaltó la importancia de las herramientas que les fueron proporcionadas en la realización de los dos ejercicios de Conteo Rápido de 2000 y 2006, durante y después de la Jornada Electoral. Herramientas como mapas de cartografía electoral y bases de datos que se actualizan constantemente fueron instrumentos informativos importantes para el análisis estadístico, pues permitieron al Comité contar con información oportuna que podía consultar constantemente.

En ese sentido, se refrendó la necesidad de disponer de los mismos mecanismos para lograr de manera exitosa el ejercicio de conteo rápido de la elección federal del próximo primero de julio de 2012.

2. **Evolución del padrón electoral.** Se enfatizó la necesidad de conocer la evolución del Padrón Electoral, así como su estado definitivo para la elección del 1ro de julio de 2012, debido a que el diseño de la muestra, necesariamente debe conocer la población o el universo que será representado, en este caso por el Padrón Electoral y el Listado Nominal. Así

como también es importante conocer las características de las personas que forman dicho padrón y la diferencia absoluta y relativa con la Lista Nominal.

3. Procesos de Reseccionamiento. Este Comité solicitó conocer la evolución del resecionamiento desde la pasada elección federal de 2006. La Dirección Ejecutiva del Registro Federal de Electores presentó la aplicación del Programa de Reseccionamiento en los años 2005, 2007, 2009 y 2010. Lo anterior en virtud de que éste procedimiento impacta en el diseño georeferencial de los estratos de la muestra en la cual se basará la estimación del Cuento Rápido.

4. Evolución del Sistema de Información de la Jornada Electoral (SIJE). Se analizó el impacto que tendrán los cambios del SIJE en la recepción de la información que recibe el Comité. Se precisó que la diferencia con respecto a 2006 será que la información estará disponible para los miembros de este Comité prácticamente en el momento en que se captura en la Junta Distrital.

Se revisaron también las medidas que se tienen previstas en caso de que se presente algún tipo de contingencia que no permita la captura de la información en algún distrito. Se precisó que es importante contar con un reporte de las casillas en las que se presentaron incidentes con la finalidad de considerar qué información no llegó a Oficinas Centrales. Se mencionó también la importancia de que se conozcan los criterios que se utilizarán para el tratamiento de dicha información en este Proceso Electoral Federal.

5. Diseño muestral. En la primera discusión sobre el diseño muestral se plantearon algunas ideas para sentar bases para la construcción de la muestra que se utilizará para el ejercicio estadístico del Cuento Rápido.

Se considera realizar un muestreo estratificado por distritos y tipo de sección urbana y no urbana. Se tendrá un muestreo aleatorio simple de casillas dentro de cada estrato con distribución proporcional. El tamaño de la muestra será de 7500, más una sobremuestra en los dos estados que tienen dos horas de diferencia con el centro del país. Este Comité seguirá trabajando en esta materia para alistar la construcción del diseño muestral a la brevedad posible.

V. Consideraciones Finales

El Comité Técnico Asesor para el Cuento Rápido continuará sesionando semanalmente, además de las reuniones que considere necesarias para acotar el rezago temporal que posee dicho Comité y así avanzar de manera oportuna en el trabajo de diseño y construcción de la muestra, definir la metodología a implementarse para la construcción del diseño muestral, establecer la metodología para la estimación de resultados y determinar los niveles de confianza y el margen de error de los estimadores.

ANEXO 2

TÉCNICO

El Comité Técnico Asesor para el Conteo Rápido definió, de acuerdo al Plan de Trabajo propuesto, diversos aspectos técnicos para la realización del mencionado ejercicio de estimación estadística. Dichas actividades se llevaron a cabo desde la instalación del órgano asesor el día 21 de marzo de 2012 y hasta la determinación de la Sala Superior del Tribunal Electoral del Poder Judicial de la Federación de revocar el acuerdo mediante el cual se constituyó el citado Comité.

Entre las actividades más relevantes destacan: el cronograma del plan de trabajo del Comité Técnico Asesor, los criterios para el diseño de la muestra, los métodos estadísticos a utilizar para la estimación de los resultados, y los procedimientos operativos para la recolección de la muestra. A continuación se describen con mayor detalle cada uno de los temas mencionados.

Plan de Trabajo

Entre los plazos cronológicos que se contemplan en el plan de trabajo del Conteo Rápido 2012, podemos mencionar los siguientes:

Se tiene planeado concluir la definición del diseño muestral y definir el procedimiento operativo para la recolección de la muestra con anterioridad al 31 de mayo.

Por otra parte, la muestra será seleccionada a más tardar el 1 de junio, actividad nodal en el proyecto. En cuanto a los simulacros, cuyos objetivos son: probar los procedimientos de reporte, captura y transmisión de la información, así como el funcionamiento de los medios de comunicación y del sistema informático, se tienen previstos para el 10 y el 24 de junio, paralelamente a los propios del Sistema de Información sobre el Desarrollo de la Jornada Electoral (SIJE).

La estimación del resultado de la votación de cada uno de los candidatos y la entrega del reporte con los resultados del ejercicio se realizará el día de la jornada electoral, es decir, el 1 de julio próximo.

El informe final que relatará el trabajo que se llevó a cabo y las recomendaciones del Comité para ejercicios de conteo rápido futuros, se entregará a finales del mismo mes de julio.

Diseño muestral

Para la definición de la muestra de casillas a considerar en la estimación estadística, se contempla utilizar un muestreo estrictamente probabilístico, el cual debe satisfacer los siguientes criterios:

1. Que todas las casillas instaladas en el país posean una probabilidad positiva y conocida de ser seleccionadas.
2. Que la selección de la muestra se realice con un mecanismo que garantice la aleatoriedad.
3. Que la muestra contenga la mayor cobertura posible.
4. Que la muestra incorpore información proveniente de todos los distritos electorales.
5. Que la muestra abarque la mayor dispersión geográfica posible, y
6. Que la muestra considere los diferentes husos horarios.

Ahora bien, entre los posibles diseños muestrales que se han considerado se encuentran los siguientes:

- a) *Muestreo aleatorio simple de casillas*. En este diseño cada casilla del universo tiene la misma probabilidad de ser seleccionada en la muestra.
- b) *Muestreo aleatorio simple de secciones y unietápico*. Se considera una muestra aleatoria simple de secciones y se capta la información de todas las casillas ubicadas en la sección seleccionada. En este caso son las secciones electorales las que tienen la misma probabilidad de ser seleccionadas en la muestra.
- c) *Muestreo estratificado simple de casillas con 300 estratos*. Se toma como estrato a cada uno de los 300 distritos electorales y se realiza un muestreo aleatorio

simple de casillas en cada uno de ellos. La distribución de la muestra a los estratos se realiza de forma proporcional al número de casillas que los conforman. Este diseño garantiza tener casillas en muestra de todos y cada uno de los 300 distritos electorales y considera que la probabilidad de selección de las casillas dentro de cada estrato es la misma.

- d) *Muestreo estratificado simple de secciones y unietápico.* Se considera a los 300 distritos electorales como estratos. En cada estrato se selecciona una muestra aleatoria simple de secciones y se capta información de todas las casillas ubicadas en la sección seleccionada. La distribución de la muestra a los estratos se realiza de forma proporcional al número de secciones que los conforman. Igual que en el inciso anterior, se garantiza tener muestra de secciones en cada distrito electoral, y la probabilidad de selección de las secciones es la misma dentro de cada estrato.
- e) *Muestreo estratificado simple de casillas con más de 300 estratos.* La estratificación se construye al considerar el cruce de los 300 distritos electorales con el tipo de sección de pertenencia de las casillas (urbana y no urbana). Este diseño es el mismo que el del inciso (c) sólo que la estratificación es más fina y como consecuencia se incrementa el número de estratos; además se garantiza tener en muestra casillas de los 300 distritos electorales y tener representado el voto en secciones urbanas y no urbanas.

De los diseños muestrales descritos, el Comité Técnico Asesor ha determinado que el **muestreo estratificado simple de casillas con más de 300 estratos** es el que cumple de mejor manera con los criterios para la selección de la muestra, además de que es el tipo de muestreo con el que se puede alcanzar una mayor precisión.

Dicho diseño muestral asegura la mayor cobertura posible al considerar los 300 distritos electorales con el tipo de sección de pertenencia de las casillas (urbana y no urbana) por lo que, además, garantiza tener representado el voto en secciones urbanas y no urbanas.

Es relevante mencionar que para el cálculo de las estimaciones con la totalidad de la muestra, se plantea una precisión de al menos medio punto porcentual (0.5%), y con un nivel de confianza de al menos 95%.

Determinación del tamaño de la muestra

En concordancia con las mejores prácticas de muestreo, el tamaño de la muestra nacional de casillas se habrá de determinar considerando los siguientes aspectos:

- Estimación de la varianza.
- Nivel de Confianza de al menos 95%.
- Precisión de las estimaciones de al menos 0.5%.

Procedimiento de selección de la muestra

Para la selección de la muestra, se partirá del marco muestral general, que considera el listado total de casillas: básicas, contiguas, extraordinarias y especiales aprobadas por los Consejos Distritales, así como las mesas de escrutinio y cómputo de la votación de los electores residentes en el extranjero.

Para cada uno de los estratos en el diseño muestral, se seleccionará un número de casillas proporcional al total de casillas en el estrato.

Esta selección se realizará mediante un programa de cómputo que asigne igual probabilidad de selección a cada una de las casillas en el estrato.

La selección de la muestra se llevará a cabo en una sesión pública en presencia de integrantes del Consejo General y ante Notario Público, de acuerdo a un Protocolo de seguridad que garantice la integridad y confidencialidad de la muestra, hasta la fecha en que deba ser utilizada.

Métodos de estimación de resultados

Para la estimación de los resultados se plantea la utilización de tres métodos estadísticos ampliamente conocidos: el Método Clásico, el Método Bayesiano y el Método Robusto.

Cada uno de ellos representa una alternativa metodológica científicamente validada y complementaria con relación a los otros dos, para llegar a resultados confiables y certeros. A continuación se describe cada uno de los métodos mencionados.

Método Clásico

El método clásico es el que se usa con más frecuencia en el muestreo. El diseño a usarse en el Conteo Rápido es un muestreo estratificado con la selección aleatoria simple dentro de cada estrato. Siendo N_h el número de casillas en cada estrato y n_h el número de casillas de la muestra.

Siendo Y_{hi} el número de votos emitidos a favor de un candidato en la casilla i del estrato h , además sea X_{hi} el número de votos totales emitidos en la casilla i del estrato h . Siendo L el número de estratos. Entonces el estimador de la proporción de votos nacional para ese partido se estima como:

$$\hat{P} = \frac{\hat{Y}}{\hat{X}} = \frac{\sum_h \hat{Y}_h}{\sum_h \hat{X}_h} = \frac{\sum_h \frac{N_h}{n_h} \sum_i Y_{hi}}{\sum_h \frac{N_h}{n_h} \sum_i X_{hi}}$$

Para tener una idea de la precisión de la estimación, es necesario calcular la varianza del estimador anterior. Esta se obtiene con las expresiones siguientes:

Primero se obtiene una varianza de una nueva variable al interior de cada estrato, esta es:

$$V_h(G_h) = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} \left[\frac{(Y_{hi} - \hat{P}X_{hi})}{\hat{X}} - \frac{1}{n_h} \sum_{i=1}^{n_h} \frac{(Y_{hi} - \hat{P}X_{hi})}{\hat{X}} \right]^2$$

$$= \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (G_{hi} - \bar{G}_h)^2$$

Donde $G_{hi} = \frac{(Y_{hi} - \hat{P}X_{hi})}{\hat{X}}$, con esta el estimador de la varianza del estimador de la proporción es

$$\hat{V}(\hat{P}) = \sum_{h=1}^L (N_h^2) \left(\frac{1}{n_h} - \frac{1}{N_h} \right) [V_h(G_h)]$$

Así, la precisión de la estimación, con una confianza de 95% es¹

$$\text{delta} = 1.96 \sqrt{\sum_{h=1}^L (N_h^2) \left(\frac{1}{n_h} - \frac{1}{N_h} \right) [V_h(G_h)]}$$

¹ Véase por ejemplo Särndal, C.E., Swensson, B., Wretman, J. (1992) "Model Assisted Survey Sampling". Springer-Verlag.

El intervalo de confianza se calcula como el estimador puntual menos y más delta. El cálculo de las varianzas se hace de forma adicional, usando el método de remuestreo Jackknife², este método será útil principalmente cuando el número de casillas cuyo resultado haya sido registrado al interior de un estrato sea menor a dos.

Método Bayesiano

El objetivo que, en términos generales, persigue la Inferencia Estadística es la descripción de un fenómeno incierto a través de un conjunto de observaciones que presentan variabilidad y constituyen sólo una fracción del total de observaciones posibles del fenómeno bajo estudio, es decir, una muestra.

En lo que se refiere a la selección de la muestra, los distintos métodos de inferencia estadística presuponen que ésta se lleva a cabo, con distintas variantes, a través de un mecanismo probabilístico. Como resultado, se propicia que los rasgos más frecuentes en la población tengan la mayor probabilidad de aparecer reflejados en la muestra.

La teoría estadística Bayesiana es relativamente reciente. Sus fundamentos fueron establecidos en la segunda mitad del siglo pasado y las técnicas de cómputo estadístico que hacen posible su aplicación, en prácticamente todo problema de inferencia, son posteriores a 1990.

El paradigma Bayesiano enfrenta todo problema de inferencia como uno de decisión en ambiente de incertidumbre y a partir de un conjunto de principios básicos determina el procedimiento general con que debe de ser resuelto.

De esta manera, establece un principio de coherencia. Todo problema debe ser resuelto con el mismo procedimiento general y los resultados producto de inferencias relativas a aspectos complementarios de un mismo fenómeno no pueden dar lugar a inconsistencias. De hecho, los principios básicos se conocen como axiomas de coherencia.

² La metodología haciendo la especificación al caso de muestreo estratificado simple, aparece en: Berger, Y.G. and Skinner, C. J. (2005). J. R. Statist. Soc, B, 79-89.

El procedimiento general de la estadística Bayesiana para la inferencia establece que:

1. Toda fuente de incertidumbre debe ser descrita a través de un modelo de probabilidad,
2. Las consecuencias que se pueden derivar de las posibles inferencias deben ser comparadas a través de una función de utilidad
3. La inferencia óptima es la que produce la utilidad esperada máxima.

En particular, el enfoque Bayesiano establece que, de la misma forma en que los datos se describen mediante un modelo de probabilidad, la incertidumbre respecto a cualquier otra cantidad o parámetro desconocido, debe describirse también un modelo de probabilidad.

Además, define un mecanismo general por medio del cual el conocimiento sobre los parámetros de interés se actualiza una vez que se tiene acceso a nuevos datos. Este mecanismo se conoce como el Teorema de Bayes³.

Para estimar el porcentaje de votos de cada candidato obtendrá como resultado en las elecciones presidenciales, se procede primero a producir la inferencia correspondiente a cada estrato y posteriormente estas inferencias se combinan tomando en cuenta los distintos tamaños de los estratos.

En cada estrato, la unidad de observación muestral es una casilla y los datos que se observan son los votos en esa casilla a favor de cada uno de los candidatos.

De esta forma, si esos votos se disponen en un vector \underline{X} , la muestra de casillas en un estrato particular da lugar a una colección de vectores.

$$\underline{X}_1, \underline{X}_2, \dots, \underline{X}_M$$

Que se consideran independientes y tales que \underline{X}_i se distribuyen según un modelo Normal multivariado con media $n_i \underline{\theta}$ y matriz de varianzas y covarianzas $n_i \underline{\Sigma}$, en donde $\underline{\theta}$ es el vector que contiene las proporciones de votos (desconocidas) a favor de los distintos candidatos en el estrato, $\underline{\Sigma}$ es una matriz (desconocida) que describe tanto la variabilidad en los votos a favor de cada candidato como la manera como estos votos

³ Una presentación exhaustiva de los fundamentos y los principales métodos del paradigma Bayesiano, se pueden encontrar en Bernardo y Smith (1994).

se relacionan entre candidatos, y finalmente, n_i corresponde con el tamaño del listado nominal de la casilla.

El objetivo es producir inferencias sobre el vector $\underline{\theta}$. Para tal fin y tomando en cuenta que tanto $\underline{\theta}$ como Σ son parámetros desconocidos, se asigna una distribución inicial de referencia, que describe la situación en que no se cuenta con información alguna antes de observar los datos de las casillas:

$$P(\underline{\theta}, \Sigma) \propto |\Sigma|^{-p/2}$$

Donde p es la dimensión de los vectores que se observan. Como consecuencia, al combinar la distribución inicial con la información provista por los datos de las casillas del estrato, se obtiene la distribución final conjunta $P(\underline{\theta}, \Sigma | \text{Datos})$, que resulta en un modelo Normal multivariada – Wishart invertida.

En particular, el vector $\underline{\theta}$ que contiene los parámetros de interés en el estrato (las proporciones de votos para cada candidato) se describe con un modelo t de Student multivariado que se caracteriza con la simulación de observaciones.

Este proceso se lleva a efecto para que cada uno de los estratos y , con las simulaciones disponibles, se obtiene una descripción de las proporciones de votos en el nivel nacional en donde cada valor simulado en este nivel se obtiene como una combinación lineal convexa de las correspondientes simulaciones en los estratos.

El resultado es una descripción, vía simulación, de la distribución conjunta de las proporciones de interés, $P(\underline{\theta} | D)$, en donde D representa la información disponible de todos los estratos. A partir de este modelo conjunto final, es posible obtener, para el candidato r -ésimo, el modelo marginal $P(\theta_r | D)$ que describe el conocimiento acumulado sobre su proporción de votos en el nivel nacional.

En particular, con este modelo marginal se pueden calcular las estimaciones puntuales y por intervalo de cada proporción. Más aún, del modelo conjunto es posible calcular algunas características que son de interés. Específicamente es posible calcular la probabilidad $P(\theta_i > \theta_j | D)$ que se interpreta como la probabilidad que se asigna, con la información disponible, al escenario en el que, al final del recuento completo, la proporción de votos para el candidato i sea mayor que la proporción de votos para el candidato j .

En resumen, el método Bayesiano permite estimar el porcentaje de votos a favor de cada uno de los candidatos y calcular la probabilidad de que el orden entre los distintos porcentajes se mantenga. La aplicación de este método se puede llevar a cabo conforme se van recibiendo las remesas de datos en la unidad concentradora de manera que se puede dar seguimiento al comportamiento de las estimaciones que produce.

Método Robusto

En este caso se considera que la muestra ha sido seleccionada de acuerdo a un esquema de Muestreo Aleatorio Simple (MAS). Bajo este supuesto, las ecuaciones que permiten estimar tanto los parámetros de interés como sus respectivos errores estándar son las más simples y, por tanto, las que más rápidamente se pueden calcular. De esta manera, con la llegada de cada remesa con nuevos reportes de resultados en las casillas de la muestra, se irán actualizando las estimaciones. A partir de estas estimaciones se irán produciendo una serie de gráficos e índices en diferentes niveles de agregación.

El MAS considera que cada elemento de la población tiene la misma probabilidad de selección. Los estimadores de los parámetros de interés son relativamente simples y esta característica ha servido para que el MAS sirva como referencia de comparación cuando se proponen esquemas más complicados que tienen como principal objetivo reducir el error estándar de las estimaciones. De esta manera, se espera que el uso de esta propuesta permita tener intervalos de confianza cuya longitud pueda ser una cota superior de los calculados en los otros dos métodos. Como en prácticamente todos los métodos de estimación estadística, es de esperar que cuando el tamaño de una muestra es grande, los estimadores tengan características similares. De esta manera, si se logra tener una gran parte de la muestra prevista, los intervalos producidos deberán ser muy parecidos a los obtenidos con los métodos Bayesiano y clásico.

En cuanto a los aspectos teóricos que fundamentan esta propuesta se tiene el siguiente modelo:

El vector de los verdaderos valores de los parámetros (P_1, \dots, P_k) es estimado del máximo verosímilmente por el vector de razones muestrales (p_1, \dots, p_k) donde $p_j = (\sum n_{ji}/n)$ con n_{ji} representando el total de votos por el partido j en la casilla i y n el total de votos registrados de la muestra. Este vector de estimadores máximo verosímiles tiene una distribución conjunta asintótica Normal multivariada con vector de medias igual a (P_1, \dots, P_k) y una matriz de varianzas y covarianzas tal que las varianzas son del tipo $(P_j(1-P_j)/n)$. Estos son resultados que, por una parte requieren grandes tamaños de muestra y también dependen de los parámetros poblacionales. Debido a que los estimadores del tipo de p_j tienen un comportamiento probabilístico que permite aproximar adecuadamente los parámetros poblacionales a partir de muestras de tamaño grande, son llamados “robustos”, este esquema de estimación se le ha denominado con tal adjetivo.

Como consecuencia, los intervalos de confianza clásicos en el MAS son del tipo

$$p_j \pm 2[(p_j(1-p_j))/n]^{1/2}$$

El proceso de incorporación de nuevas remesas permite que los intervalos de confianza vayan teniendo longitud monótonamente no creciente. Esta característica implica que conforme transcurra el arribo de información se tendrán mejores estimaciones y, en el límite del 100% de la muestra, tener calidades comparables a los otros dos procesos que se estarán realizando simultáneamente.

Procedimiento operativo para la recolección de la información muestral

Para la recopilación de los datos, los presidentes de las mesas directivas, de las casillas seleccionadas, darán acceso a los Capacitadores Asistentes y Supervisores Electorales a la información plasmada en las actas de escrutinio y cómputo de la elección de Presidente, inmediatamente después de haber efectuado el conteo de los votos y el llenado del acta mencionada.

Los Capacitadores Asistentes y Supervisores Electorales transmitirán, mediante los medios de comunicación disponibles, los resultados de la casilla en cuestión a su respectiva sede distrital, en donde un par de capturistas integrarán la información recibida al sistema informático desarrollado para tal fin y que operará a través de la red informática del IFE (RedIFE) bajo un esquema de estricta seguridad. Este mecanismo

permitirá que el Comité Técnico Asesor reciba remesas de información íntegra y confiable.

Los miembros del Comité realizarán los ejercicios de estimación con los métodos definidos (Bayesiano, Robusto y Clásico). Con dicha información se llevarán a cabo los análisis necesarios y se entregará un informe al Presidente del Consejo y al Consejo General para que se comuniquen las tendencias electorales la noche del día de la jornada electoral.

Para la correcta operación de todo el aparato logístico de recolección de la información de las casillas que forman parte de la muestra, se tiene previsto realizar dos simulacros: los días 10 y 24 de junio de 2012. Estos ejercicios se realizarán al concluir los simulacros del Sistema de Información sobre el Desarrollo de la Jornada Electoral (SIJE) y serán en día domingo buscando emular al máximo las condiciones que prevalecerán el día de la jornada electoral.

El objetivo de estos ejercicios es probar los procedimientos de reporte, captura y transmisión de la información, así como el funcionamiento de los medios de comunicación y del sistema informático, para detectar oportunamente cualquier posible contingencia en los aspectos técnico-logísticos, con el fin de realizar los ajustes necesarios y así garantizar el envío oportuno y la recepción de los datos de la votación en las casillas de la muestra de tal manera que el Comité Técnico Asesor para el Conteo Rápido, cuente con la información requerida el día de la jornada electoral.