

Comité Técnico Asesor para el Conteo Rápido de la Elección Ordinaria para Gobernador de Oaxaca

Criterios científicos, logísticos y operativos para
la realización del Conteo Rápido y protocolo
para selección de muestra

Mayo, 2016

Contenido

I. Antecedentes	3
II. Criterios científicos	4
2.1 Esquema de muestreo	4
III. Selección y resguardo de la muestra	10
3.1 Consideraciones generales para la selección de la muestra	10
3.2 Protocolo de selección y resguardo de la muestra	10

I. Antecedentes

El 30 de abril de 2016, el Consejo General del Instituto Estatal Electoral y de Participación Ciudadana de Oaxaca aprobó, mediante acuerdo IEEPCO-CG-67/2016, la realización del Conteo Rápido en Elección Ordinaria de Gobernador en el estado de Oaxaca con el fin de conocer las tendencias de los resultados de la votación el día de la jornada electoral. También aprobó la integración del Comité Técnico Asesor en la materia (COTECORA).

En apego a las disposiciones normativas en la materia, especialmente los LINEAMIENTOS GENERALES DEL INSTITUTO NACIONAL ELECTORAL PARA EL DISEÑO, IMPLEMENTACIÓN Y OPERACIÓN DE LOS CONTEOS RÁPIDOS DE CARÁCTER INSTITUCIONAL EN LOS PROCESOS ELECTORALES LOCALES 2015-2016 y 2016-2017 emitidos el 30 de octubre de 2015 mediante el acuerdo INE/CG921/2015, el COTECORA elaboró los criterios científicos. Además, conoció el diseño de los procedimientos logísticos y operativos para el acopio de los resultados de las casillas, así como el protocolo para la selección de la muestra, y los consideró pertinentes para los fines del Conteo Rápido. Este documento describe los criterios científicos, logísticos y operativos así como el protocolo de selección de la muestra.

II. Criterios científicos

Los criterios científicos del Conteo Rápido de la elección ordinaria de Oaxaca 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 y el porcentaje de ciudadanos que acudieron a votar.

El enfoque estadístico que se usará para la estimación de los resultados de la elección es el Clásico (algunas veces llamado Frecuentista). Cabe mencionar que, dentro de este enfoque, es posible incorporar distintos matices metodológicos.

Se aplicarán diversos procedimientos de estimación para obtener los intervalos de confianza de los resultados del Conteo Rápido de la elección de gobernador. No se esperan resultados idénticos de los distintos procedimientos porque los métodos que se utilizarán introducen elementos aleatorios. Las estimaciones están basadas en una muestra. Sin embargo, el objetivo de los procedimientos de estimación será el mismo, y las estimaciones se consolidarán para presentar resultados únicos en forma de intervalos de confianza referidos al porcentaje de votos a favor de cada candidato a la gubernatura de Oaxaca.

2.1 Esquema de muestreo

Diseño de muestreo

Las estrategias de estimación que implementarán los miembros del COTECORA estarán basadas en la misma información muestral. El proceso de selección será un muestreo aleatorio polietápico, con una primera etapa donde se seleccionarán secciones electorales dentro de cada distrito. La segunda etapa

de selección trata a las secciones como conglomerados. Se recopilará información de las casillas dentro de las secciones seleccionadas. La Tabla 1 describe este procedimiento, en términos de los distritos, y los números de secciones electorales y casillas esperado en la muestra de cada uno.

Tabla 1. Tamaño de muestra y distribución esperada para las elecciones de gobernador del estado de Oaxaca

Número de distrito	Cabecera Distrital	Número de secciones en el Distrito	Número de secciones en el Distrito esperadas en la muestra	Número de Casillas en el Distrito	Número promedio de casillas por sección en el distrito	Número de casillas esperadas en la muestra
1	ACATLÁN DE PÉREZ FIGUEROA	127	21	211	1.7	34
2	SAN JUAN BAUTISTA TUXTEPEC	113	18	187	1.7	31
3	LOMA BONITA	103	17	197	1.9	32
4	TEOTITLÁN DE FLORES MAGÓN	81	13	212	2.6	35
5	ASUNCIÓN NOCHIXTLÁN	157	26	249	1.6	41
6	HEROICA CIUDAD DE HUAJUAPAN DE LEÓN	86	14	251	2.9	41
7	PUTLA VILLA DE GUERRERO	166	27	216	1.3	35
8	HEROICA CIUDAD DE TLAXIACO	91	15	229	2.5	37
9	IXTLÁN DE JUÁREZ	87	14	223	2.6	36
10	SAN PEDRO Y SAN PABLO AYUTLA	109	18	179	1.6	29
11	MATÍAS ROMERO AVENDAÑO	121	20	234	1.9	38
12	SANTA LUCIA DEL CAMINO	128	21	176	1.4	29
13	OAXACA DE JUÁREZ (SUR)	92	15	206	2.2	34
14	OAXACA DE JUÁREZ (NORTE)	102	17	211	2.1	34
15	SANTA CRUZ XOXOCOTLÁN	119	19	176	1.5	29
16	ZIMATLÁN DE ÁLVAREZ	39	6	207	5.3	34
17	TLACOLULA DE MATAMOROS	89	15	234	2.6	38

18	SANTO DOMINGO TEHUANTEPEC	61	10	187	3.1	31
19	SALINA CRUZ	70	11	208	3.0	34
20	HEROICA CIUDAD DE JUCHITÁN DE ZARAGOZA	122	20	199	1.6	32
21	HEROICA CIUDAD DE EJUTLA DE CRESPO	108	18	176	1.6	29
22	SANTIAGO PINOTEPA NACIONAL	64	10	235	3.7	38
23	SAN PEDRO MIXTEPEC	62	10	186	3.0	30
24	MAHUATLÁN DE PORFIRIO DÍAZ	75	12	204	2.7	33
25	SAN PEDRO POCHUTLA	78	13	188	2.4	31
	TOTALES	2450	400	5181	2.1	846

La muestra total será de 450 secciones. Se espera recibir información de al menos 400 de ellas. Ante la posibilidad de no respuesta, se seleccionará una muestra adicional de 50 secciones electorales para sustitución, dos por cada Distrito Local. El número de casillas en la muestra final dependerá de las secciones seleccionadas. Se espera tener 846 casillas, asumiendo que el número de casillas en la muestra de cada distrito es igual al número de secciones por el promedio de casillas por sección en cada distrito. Dado que el número de casillas por sección no es constante, este número puede variar dependiendo de las secciones seleccionadas. La distribución de secciones en cada estrato se realizará vía afijación proporcional u óptima. Para la elaboración de la Tabla 1 se utiliza como referencia la distribución de casillas instaladas para la elección local del 5 junio de 2016.

Consideraciones para definir el tamaño de muestra y diseño de muestreo

Ideas básicas del muestreo probabilístico para el Conteo Rápido de Oaxaca

El objetivo del Conteo Rápido es estimar la proporción p de votos en favor de un candidato a gobernador¹. Entonces, se selecciona una muestra aleatoria de n casillas de un total de N y con la información recuperada se calcula el estimador \hat{p} . Mediante el muestreo probabilístico es posible definir **estrategias de selección, tamaños de muestra n y estimadores** para asegurar que

$$|p - \hat{p}| \leq d, \quad (1)$$

con un 95% de confianza. A la cantidad d comúnmente se le conoce como precisión o error máximo aceptable en la estimación. La expresión (1) se puede escribir de forma equivalente en términos de intervalos de confianza

$$\hat{p} - d \leq p \leq \hat{p} + d. \quad (2)$$

Esto significa que si, por ejemplo, se extraen 100,000 muestras distintas e independientes una de otra (cada una siguiendo la misma estrategia de selección, usando el mismo tamaño de muestra y el mismo estimador) y con cada muestra se hace una estimación

$$\hat{p}_1, \hat{p}_2, \hat{p}_3, \hat{p}_4, \hat{p}_5, \dots, \hat{p}_{100,000}$$

al menos 95,000 estimaciones cumplirán con la condición (1).

Precisión en la estimación

Lo anterior brinda un procedimiento mediante simulación para obtener las precisiones de varias estrategias de selección y estimadores usando diferentes tamaños de muestra y así decidir cuál es el que mejor se ajusta a los objetivos del Conteo Rápido.

¹ Esto se hará para cada contendiente así como para la estimación de la votación en favor de candidatos no registrados y votación nula, sin embargo para efectos de la explicación es mejor concentrarse en la estimación para un solo contendiente.

Con cada una de las 100,000 muestras simuladas se calculó el valor del estimador de razón:

$$\hat{p}_i = \frac{\sum_{h=1}^H \sum_{i=1}^I N_{hi} \bar{y}_{hij}}{\sum_{h=1}^H \sum_{i=1}^I \sum_{j=1}^J N_{hi} \bar{y}_{hij}} \quad (3)$$

En donde j es el candidato j -ésimo de J candidatos (incluyendo también el número de votos por candidatos no registrados y nulos), h es el número de estrato (distrito), H el número total de estratos, i es el número de conglomerado (sección) en el estrato (distrito) h , I el número total de conglomerados (secciones) en el distrito h . Se debe notar que I varía de distrito a distrito. N_{hi} es el número de casillas instaladas en el estrato (distrito) h , conglomerado (sección) i y \bar{y}_{hij} es el número promedio de votos en las casillas seleccionadas en muestra para el candidato a gobernador por el partido i en el estrato h , conglomerado i .

La precisión d implica que 2.5% de las estimaciones del porcentaje de votos de un candidato j quedarán por debajo del valor $\hat{p} - d$, y que 2.5% de las estimaciones quedarán por arriba del valor $\hat{p} + d$. Entre ambos límites $[\hat{p} - d \text{ y } \hat{p} + d]$ se encuentra el 95% de las estimaciones. Entonces, regresando a las expresión (2) tenemos que

$$q_{0.025} = \hat{p} - d \text{ y } q_{0.975} = \hat{p} + d.$$

Y la precisión se obtendría mediante

$$d = \max\{p - q_{0.025}, q_{0.975} - p\}.$$

Así, utilizando una simulación de 100,000 muestras de 400 secciones, para los resultados de diputado local de 2013, se obtiene una precisión de $d = 1.16\%$.

Siguiendo el mismo procedimiento se obtuvieron las precisiones para varias estrategias de selección, tamaños de muestra y estimadores.

Población de referencia para fijar el tamaño de muestra

En todos los ejercicios realizados se tomaron como referencia los cómputos distritales de la elección para diputados en la elección ordinaria de 2013 en Oaxaca.

Estratificaciones

Las estratificaciones consideradas en los ejercicios de simulación fueron 25 distritos locales del estado de Oaxaca.

Por último, es importante considerar que la precisión de los resultados depende tanto en la etapa de planeación del ejercicio como de 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 se fija este tamaño de muestra tomando como referencia los resultados de una elección previa y que un porcentaje de no respuesta que no conocemos, la precisión de la muestra no se conoce sino hasta después de realizado el ejercicio. Un segundo elemento que influye en la precisión final de los resultados, es qué métodos de estimación se utilizan para producir las estimaciones de los porcentajes de votación para cada candidato. Una vez concluida la elección y con toda la información disponible se determina la precisión realmente obtenida.

III. Selección y resguardo de la muestra

3.1 Consideraciones generales para la selección de la muestra

La muestra con la que se estimará la votación a favor de los candidatos a la gubernatura de Oaxaca se realizará en un acto público el sábado 14 de mayo de 2016. En el acto estará presente un notario público que dará fe del desarrollo del protocolo desde la instalación del software requerido hasta la obtención y resguardo de la muestra definitiva.

Para seleccionar la muestra se usará de un equipo de cómputo habilitado con software estadístico. El marco muestral será el listado de las secciones y las casillas pertenecientes a éstas aprobado para las elecciones y la selección se realizará de acuerdo con el diseño de muestreo establecido.²

3.2 Protocolo de selección y resguardo de la muestra

Para la selección y resguardo de la muestra se realizarán las siguientes actividades:

Instalación

1. Personal de la IEEPCO realizará ante notario público la validación del equipo de cómputo donde se instalará el programa para la obtención de la muestra.
2. El COTECORA entregará el código necesario para la selección de la muestra.

² Se utilizará la última lista oficial de secciones y casillas que se tenga en el IEEPCO.

3. A la vista de los asistentes se introducirá el código que se utilizará.
4. Se solicitará a un miembro del COTECORA iniciar la etapa de ejecución de la selección de la muestra.

Selección de la muestra

1. Un miembro del COTECORA ejecutará el programa para seleccionar la muestra.

Resguardo de la muestra

1. La muestra será grabada en tres discos compactos no regrabables.
2. El archivo será eliminado de la máquina.
3. Un disco será entregado al Secretario Técnico con el fin de que se preparen los trabajos de acopio de datos.
4. Los otros dos discos con la muestra serán guardado en un sobre el cuál será sellado y rubricado por el Notario Público.
5. El Notario Público entregará a la autoridad del IEEPCO designada los sobres a los que se refiere el numeral anterior para su resguardo.