

MESA 3



CONTEO RÁPIDO IFE



RELATORÍA

Roy Campos

PONENCIAS

Raúl Rueda

Rubén Hernández

Edmundo Berumen

RELATORÍA

.....
Roy Campos

El objetivo de esta mesa fue presentar la metodología, el entorno y las circunstancias en las que se realizó el conteo rápido que el Instituto Federal Electoral (IFE) encomendó levantar a tres empresas encuestadoras.

El desarrollo de las discusiones y las presentaciones dio oportunidad a incorporar a la mesa temas como el del papel de la autoridad como contratante de conteos rápidos, la forma de anunciarlos y las diferencias que se presentan en este tipo de ejercicios muestrales, según el contratante.

En la exposición quedó muy claro que este conteo rápido tuvo una dimensión política que no poseían otros conteos, y que esa dimensión influyó en el diseño muestral y en la decisión de hacerlo en tres réplicas con distintas empresas.

El comité técnico que se formó al interior del IFE para diseñar este conteo rápido decidió:

- ✓ Un diseño muestral que permitiera generar tres resultados confiables, pero que a su vez permitiera agregar los tres ejercicios en un resultado que diera más certidumbre al anuncio del IFE.
- ✓ El diseño de muestra consistió básicamente en dividir al país en los 300 distritos electorales federales y dividir dos mil 550 secciones en forma proporcional en ellos, seleccionándolas con igual probabilidad y dividiendo después la muestra en tres réplicas de 850 con características similares, asignando aleatoriamente una a cada empresa.

La decisión de utilizar 300 estratos fue explicada como política, ya que algunos consejeros pedían una distribución de la muestra con esa garantía.

Para la estimación de resultados y precisiones cada empresa utilizó el paquete *Clusters* y el comité técnico decidió, además, construir un modelo que permitiera tener mayor certeza al anunciar al ganador, para lo cual utilizaron elementos georeferenciales para analizar la completez y composición de la muestra y una prueba de hipótesis que medía la probabilidad de equivocarse en el ganador, la que estimaron en menos de un milésimo el 2 de julio a la hora del anuncio.

Las diferencias que se ubicaron entre un conteo rápido contratado por la autoridad electoral y uno contratando por otro tipo de contratante fueron:

- ✓ El costo político de un error es mucho mayor.
- ✓ El diseño estadístico se ve influenciado por los consejeros electorales.

- ✓ En el anuncio de resultados se reduce la importancia de la “oportunidad”.
- ✓ El anuncio debe incluir los votos nulos y no registrados para coincidir con los datos finales.

Hubo ocho manifestaciones sobre la conveniencia de que los métodos de estimación respeten los parámetros del diseño de muestreo utilizado. ■

PONENCIAS MESA 3



Raúl Rueda
IFE

IIMAS-UNAM E IFE

Se presenta una descripción de los trabajos que fueron encomendados al Consejo Técnico por el Instituto Federal Electoral (IFE), relativos a los conteos rápidos realizados el día 2 de julio de 2000. La descripción es no técnica, haciendo énfasis en los problemas relevantes y mencionando la manera en que fueron resueltos

1. Antecedentes

Por acuerdo del Consejo General, el IFE decidió hacer conteos rápidos para estimar la proporción de votos que recibiría cada partido/coalición en la votación presidencial del día 2 de julio de 2000. Como grupo técnico asesor, el Consejo General creó un Consejo Técnico (CT) que estuvo formado por Víctor Guerra (coordinador) (DGSCA, UNAM), Rubén Hernández (ITAM), Ignacio Méndez (IIMAS, UNAM), Raúl Rueda (IIMAS, UNAM) y Gabriel Vera (Banco de México). Este grupo tuvo la colaboración, para la realización de sus trabajos, de Analía Babinsky, Hortensia Moreno, Federico O'Reilly, Salvador Zamora, Martha Zertuche y personal del Registro Federal de Electores.

El Consejo General decidió que se contrataran tres empresas con experiencia en el ramo y que los esfuerzos del Comité Técnico se abocaran principalmente a la solución de tres problemas: planteamiento de un diseño muestral, coordinación del trabajo con las empresas y la consolidación de los resultados obtenidos por cada una de ellas. Finalmente, el Consejo General decidió que el Consejero Presidente debería hacer un anuncio sobre los resultados de los conteos rápidos a las 23:00 horas del 2 de julio, lo que acotaba el trabajo del Comité. Las empresas que se contrataron fueron, en orden alfabético, Alduncin y Asociados, Berumen y Gallup México.

El objetivo fue dar una descripción no técnica de los trabajos efectuados por el Comité Técnico el día de la elección, los problemas estadísticos que surgieron en estos ejercicios y las decisiones que se tomaron, por lo que la estructura de este trabajo es como sigue: en la siguiente sección se justificará el uso de la estadística al realizar un conteo rápido. Las tres secciones siguientes describirán el diseño de la muestra, el trabajo con las empresas y la consolidación de los resultados, respectivamente. La sexta sección es sobre lo que el Consejero Presidente dijo sobre los conteos, y el trabajo que falta por hacer se comenta en la última sección.

2. El problema estadístico

El objetivo de todo conteo rápido es dar una estimación oportuna de las proporciones de votación a favor de cada uno de los participantes en una elección, con base en una muestra de secciones o casillas. Como existe incertidumbre sobre el resultado de la elección, la única manera científica de cuantificar esta incertidumbre es mediante una medida de probabilidad. Esto significa que deberá elegirse una *muestra aleatoria* de secciones, y con los resultados electorales de esta muestra inferir sobre los resultados nacionales. La única característica que se le pide a una muestra para que sea *aleatoria* es que cada sección tenga una probabilidad conocida de pertenecer a la muestra.

En este caso la población bajo estudio puede considerarse conocida: el total de lista nominal reportada en todas las secciones del país. Sin embargo, la población real es el conjunto de personas que votan el día de la elección, por lo que un análisis condicional al total de votos emitidos es recomendable. Por otro lado, las proporciones que el Programa de Resultados Electorales Preliminares (PREP) reporta durante la noche de la elección es referido a votos emitidos. Por estas dos razones se decidió que el objetivo era estimar las proporciones nacionales de votación con respecto al total de votos emitidos.

El IFE, como la autoridad electoral, debería estar seguro sobre los resultados que anunciaría, pues el costo político de hacer un mal anuncio era gigantesco, por lo que era preferible tener poca precisión en el pronóstico, pero estar seguro de anunciar correctamente al ganador.

3. Diseño de la muestra

Se pidió diseñar una muestra que incluyera secciones de todos los distritos electorales del país y que los ejercicios de las empresas fueran independientes entre sí, de manera que el proceso resultara robusto. Como el CT debía consolidar los resultados de las tres empresas, se pensaron diferentes métodos de estimación para combinar los resultados, lo que llevó al CT a proponer el siguiente diseño muestral:

Se consideró a los 300 distritos electorales del país como estratos y dentro de cada uno se elegirían, proporcionalmente, al número de secciones en cada distrito, secciones electorales que serían censadas el día de la elección. El tamaño de la muestra se fijó en $3n$. Posteriormente, la muestra se dividió aleatoriamente, manteniendo la proporcionalidad en tres submuestras de tamaño n y cada una de ellas asignada aleatoriamente a cada una de las empresas participantes.

Así, cada submuestra es, en realidad, una muestra aleatoria de tamaño n distribuida proporcionalmente

al número de secciones en cada distrito con los distritos, formando una estratificación geográfica del país

Este diseño fue discutido con las empresas participantes y aceptado por ellas antes de firmar los respectivos contratos de prestación de servicios. Hubo dudas sobre el diseño muestral entre ellas, si la proporción urbana-rural de las secciones estaría presente en la muestra. La elección totalmente al azar, garantiza –con probabilidad muy alta– que esta proporción sería respetada, tanto en el número de secciones, como en el total de población con derecho a votar en cada clasificación. Al final, cuando las muestras fueron obtenidas, esto quedó confirmado.

Una razón muy importante para hacer así el diseño tenía que ver con el problema de combinar los resultados de las tres empresas, específicamente con el tipo de procedimientos de estimación que se pensaban usar.

Conjuntamente con las empresas se decidió que $n = 850$. Este tamaño muestral tomaba en cuenta el efecto de diseño por no respuesta, que permitiría a cada empresa obtener estimaciones con la precisión y confianza establecidas en la licitación emitida por el IFE.

La muestra fue obtenida por el CT en presencia del presidente del Consejo General del IFE, consejeros electorales y representantes de partidos políticos, manteniendo su confidencialidad.

4. Coordinación con las empresas

En una primera plática con las empresas se les informó del diseño muestral y se les especificó qué es lo que se esperaba de ellas. Concretamente, se les dijo que tendrían que enviar con cierta regularidad y a partir de cierta hora, los archivos que usarían para hacer sus estimaciones y las propias estimaciones con un comentario sobre si eran o no definitivas y que el CT sería el encargado de consolidar toda la información, respetando lo que cada empresa reportara. En las reuniones siguientes se trataron diferentes asuntos: las empresas especificaron los métodos de inferencia que usarían para hacer sus estimaciones y cómo las generarían; se quedó de acuerdo en los formatos de las bases que enviarían con la información y, sobre todo, las necesidades de las empresas, como equipo de cómputo, comunicaciones, mapas de secciones, información sobre listas nominales y número de casillas y secciones, entre otras cuestiones.

Además de decidir el tamaño de muestra, los acuerdos más relevantes fueron dos: el primero, que las empresas usarían el paquete *Clusters* para hacer sus estimaciones, copia del cual fue entregado al CT, sobre todo para tener muy claros los métodos de estimación que se usarían. El segundo acuerdo fue respecto a la frecuencia y for-

mato en el que entregarían los resultados. Con respecto a la primera, se quedó que la información con sus resultados y una copia de la base de datos, fuera entregada con una frecuencia de media hora a partir de las 19:00 horas y que sería de 15 minutos a partir de las 21:00 horas. El formato especificaba que se darían los intervalos de confianza para las primeras tres fuerzas principales, así como el estimador puntual, la hora de entrega, el porcentaje de casillas levantado y el tipo de estratificación que se consideraba. Además, se incluiría un comentario sobre la posición de la empresa respecto a los resultados entregados.

La actitud de las empresas fue de cooperación y siempre propositiva para resolver todos los problemas que se iban presentando. Las tres empresas estuvieron dispuestas, en todo momento, a cumplir los requerimientos que el CT le solicitaba, de manera eficiente y profesional. Debe mencionarse además que el tiempo para que planearan todo su trabajo fue muy corto, en especial lo relativo al operativo de campo.

Un problema que surgió, afortunadamente con poca frecuencia, fue que algunas secciones de las elegidas en la muestra, serían imposibles de visitar el día de la elección, ya sea por problemas de comunicación, por problemas políticos (zonas ocupadas) o por otros problemas (poca seguridad para los encuestadores). La decisión del CT, en el sentido estadístico más purista, fue la de suprimir estas secciones y no buscar sustituciones, para no alterar las probabilidades de selección. Esto implicaba que habría que considerar un elemento de censura en la información recibida y tratar de incorporarlo al método de estimación o bien, considerar que la eliminación de estas secciones no alteraban considerablemente las probabilidades de selección del resto de las secciones. Dado que fueron muy pocas las secciones que se eliminaron, se eligió la segunda opción.

5. Consolidación de resultados

Puede decirse que éste fue el problema técnico (estadístico) más importante al que se enfrentó el CT. Cada empresa era libre de decidir si las estimaciones que enviaba eran definitivas. Esto dependía no sólo si la cantidad de secciones pertenecientes a la muestra que hubieran recibido permitía tomar una decisión, sino también si se podía decir que esa parte de la muestra, en efecto, podía considerarse una muestra aleatoria en sí misma. Esto significa que la información muestral -incompleta- recibida no contuviera sesgos que pudieran afectar las propias estimaciones.

Para la consolidación de los resultados el CT tenía, básicamente, dos posibilidades: combinar de manera

lineal los resultados de las tres empresas o juntar toda la información y hacer el análisis como si fuera una sola muestra. El diseño muestral establecido permitía cualquiera de los dos caminos. Por otro lado, se tenían diferentes opciones para realizar las estimaciones: usar, al igual que las empresas, el paquete *Clusters*; programar directamente las expresiones que *Clusters* usa, con el fin de hacer un doble chequeo y, por último, utilizar otra forma de estimar que considerara que la muestra era sólo parte de una muestra más grande. En cualquier caso, era importante determinar si la información recibida podía ser considerada en sí misma una muestra aleatoria de secciones.

Primero se decidió que la información provista por cada empresa debería tener el mismo peso y, por lo tanto se deberían de juntar las tres últimas bases proporcionadas por las empresas y después de pasar por filtros que limpiaran la información de posibles errores de captura, construir una base considerando a toda la información como una sola muestra. Esta es una de las razones de por qué los resultados anunciados por el Consejero Presidente, como responsabilidad del CT, no son un promedio de los emitidos por las empresas, además del hecho de que el CT usó información de las tres empresas a pesar de que se anunció que sólo utilizó la de las dos empresas que presentaron resultados.

Evaluar si las secciones que han llegado hasta cierto momento son una muestra aleatoria, no es posible estadísticamente. Sin embargo, diferentes aproximaciones pueden dar cierta confianza de que la información recibida no presenta sesgos y que lo que se anunciara tiene cierta probabilidad de estar bien. El camino que eligió el CT fue la construcción de un sistema automático de "escenarios" que consistía someramente en lo siguiente: Se hacían las estimaciones usando la información disponible, reportando el porcentaje de muestra recibido. Con esta información, el peor escenario que podría ocurrir con el resto de la información era que ésta viniera totalmente a favor de uno de los partidos. Se podían plantear diferentes escenarios, y para cada uno de ellos se reestimaban las proporciones de cada partido. Si el resultado de estas estimaciones no cambiaba significativamente en un conjunto de escenarios razonables, se decidía que la información disponible no tenía sesgos y se consideraba que formaba una muestra aleatoria. Este análisis se apoyaba en un sistema georeferencial que personal del Registro Federal de Electores había desarrollado y que, a petición del CT, graficaba la distribución geográfica de la muestra que había llegado y, muy importante también, la que faltaba por llegar. Este despliegue gráfico se hacía para cada compañía y conjuntamente,

a nivel seccional, distrital, estatal o nacional, y tomaba en cuenta la clasificación urbano-rural de las secciones. Además de dar información sobre la distribución geográfica de la información recibida, se tenía información sobre la cobertura de cada empresa. Esto permitía hacer una evaluación, subjetiva, de la existencia de posibles sesgos en la información recibida.

Ya decidido que la información con la que se contaba podía ser considerada una muestra, la consolidación de los resultados se hizo de tres maneras distintas, usando, en los tres casos, una sola base de datos. La primera consistió en obtener las estimaciones usando el paquete *Clusters*. El objetivo, al hacer esto, era tener compatibilidad con los métodos utilizados por las empresas y hacer unas verificaciones rápidas. La segunda aproximación se basó en un programa escrito específicamente para este problema, basado en el programa JMP –una versión portable, interactiva y gráfica del paquete SAS–, que usaba los estimadores utilizados por el paquete *Clusters*, con ligeras modificaciones. Finalmente, el tercer enfoque consistió en el desarrollo de un modelo, basado en el enfoque bayesiano de la estadística y que tiene fuertes conexiones con la estadística fiducial, completamente distinto a los dos anteriores. Hay dos aspectos relevantes en este análisis que difieren de los dos anteriores. Primero, las estimaciones se hacen condicionales a la información recibida y no a la información que se esperaría recibir, lo que hace que los estimadores sean más realistas y, segundo, puede calcularse fácilmente la distribución (final o fiducial) de la diferencia de votación entre dos fuerzas electorales condicional al total de votos válidos, dando así información sobre la probabilidad (y no confianza) de cometer un error al decidir cuál partido lleva la delantera. Aunque esta distribución puede calcularse de manera exacta, no tiene expresión analítica, así que se utilizó una aproximación normal. La solución numérica exacta resulta ser, básicamente, la misma.

6. ¿Qué se dijo?

Cada empresa era responsable de los resultados que producía, e incluso tenía la libertad de decidir “no

salir”, como ocurrió con una de ellas. Esta información fue dada a conocer por el Consejero Presidente a las 23:00 horas, como se había anunciado. Además de los resultados de las tres empresas, el Consejero Presidente anunció los resultados que el CT le presentó y que reflejaban la consolidación de la información. La decisión del CT, dada la gravedad de las consecuencias de una “mala salida”, fue la de perder precisión en las estimaciones, pero tener seguridad en el orden que se daba. Los números que se dieron al Consejero Presidente aseguraban una probabilidad (fiducial o final) por abajo de uno sobre el millar de error en el orden y fueron construidos basándose en los resultados de los tres ejercicios descritos en la sección anterior y con la información que las tres empresas habían enviado al CT a las 22:00 horas aproximadamente.

7. ¿Qué falta hacer?

Una ventaja de la manera en que se diseñó todo, es que es totalmente replicable. Se tienen todos los archivos –muestra, archivos de datos, programas de estimación y mapas, entre otros– que permiten repetir todo el proceso casi de manera puntual.

El Consejo Técnico deberá elaborar un reporte de todas las actividades que desarrolló, en donde sería recomendable que se auditara el trabajo realizado por todos los actores involucrados, con la finalidad de corregir errores e ir tendiendo a un sistema automatizado que permita tomar decisiones en poco tiempo.

8. Agradecimientos

Al Comité Organizador del taller, por la invitación a participar y a algunos de los participantes cuyos comentarios y observaciones a la exposición, mejoraron el contenido de este trabajo. Al IFE por la invitación a participar en el Consejo Técnico y al resto del Consejo Técnico por el intercambio de ideas. Este trabajo no refleja la posición del Consejo Técnico ni de las empresas participantes, es sólo una apreciación personal sobre los trabajos realizados y el autor es el único responsable de lo aquí presentado. ■

PONENCIAS MESA 3

.....

Rubén Hernández

IFE

Bueno, simplemente para cerrar esta mesa quisiera yo hacer algunos comentarios. Algo se mencionaba en la Mesa 2. Manuel Barberena había hablado de la enorme importancia que ha tenido este ejercicio en esta elección de 2000. Yo creo que esta situación se sintetizaría en esa famosa frase que José Woldenberg mencionó en cadena nacional a las 23:00 horas y que se resumiría en este pensamiento, de que hay suficiente evidencia estadística.

Esto, en verdad, yo creo que es un cambio cualitativo enorme el hecho de que ya nuestro país esté alcanzando alguno de los patrones que son comunes en otros países, pero que en el nuestro se estén dando por buenos los resultados de una elección con base en un ejercicio estrictamente estadístico, nos parece muy relevante.

Creo que lo hemos alcanzado, no ha sido simple. Simplemente también quiero recordarles cuál era el contexto en 1994. En 1994 nuestro gran problema ni siquiera eran los conteos o cualquier cosa, era el padrón electoral en un primer momento. El país, no se lo tengo que describir, todo mundo lo recordamos. Creo que en 1997 fue esa transición donde la estadística y todos estos procesos que van acompañándola, como instrumento, se consolidaron.

Creo pues, sinceramente que a partir de ahora no es ya una cuestión de estar convenciendo gente de si la estadística estará o no en el punto central del asunto.

Bueno, tal vez Enrique Alduncin habló de un tema que no era muy lejano finalmente del asunto. Creo que cuando se es profeta es muy fácil hacer pronósticos, lo malo es que se equivocan a veces y más que lo que se esperaba.

Creo que no aceptamos la apuesta que se hacía por ahí, pero hubiera perdido este señor, lo malo es que me llamó Raúl Rueda y eso sí. Bueno, fuera de ahí es cierto, es que ser profeta y hacer pronóstico, pues es muy fácil.

Lo que quiero hacer en esta intervención es un brevísimo recuento de cómo fue la integración. Evidentemente quedan muchos puntos por ahí, los mencionaba ya Edmundo Berumen, algunos de carácter muy complicado. Creo que la idea fue que el CT cayó, por así decirlo, en esta situación que era totalmente inédita en nuestro sistema; no seleccionó a las compañías, no tuvo a su alcance todos los recursos que tal vez hubiese podido hacer porque sinceramente creo que el tiempo fue absolutamente en nuestra contra, no teníamos ningún grado de libertad.

Se menciona que tal vez el diseño se hubiera podido mejorar, por supuesto. Ayer partimos de un principio de que por definición nunca nos va a gustar la encuesta de nuestro vecino y menos la nuestra, estamos de acuerdo, Edmundo Berumen.

Por lo demás, estaba yo mencionando aquí la salida de las 22:00 horas que fue con la que José Woldenberg armó el comunicado en cadena nacional y podemos ver que prácticamente la compañía de Enrique Alduncin no había entregado; sí se procesó la información que se produjo y en los estimadores, en los intervalos que se mencionaron en el comunicado fueron incluidos, por supuesto que sí.

Bueno, lo que quería yo más que nada es centrarme ahora, simplemente en las tareas que se tuvieron aquí en este asunto a la hora de la integración de los resultados; esas dos horas fueron muy complicadas, hay que imaginar la situación que teníamos, tal vez no sea muy relevante para algunos, lo que queremos es mostrar cómo fue eso.

Hemos venido repitiendo en diferentes foros que queríamos que nuestros colegas, compañeros de otras entidades, compartan sus datos con nosotros. Se mencionó ayer acerca de que deberían estar a disposición de todo mundo esas bases de datos.

Lo que yo quiero mostrarles hoy es un producto que el Registro Federal de Electores junto con el equipo que hemos estado trabajando en este conteo rápido, hemos querido compartir esa experiencia con ustedes.

Entonces, es un sistema que nos va contando prácticamente, minuto por minuto, cómo sucedió el proceso. La idea también es ponerlo a la disposición de todos ustedes y que puedan replicarlo. Están las fórmulas, está todo. De tal manera que este es un sistema que van a tener ustedes dentro de muy poco en la página web del IFE.

Contamos, pues, con un menú típico, donde tenemos una serie de temas, modelo matemático; esa es una de las actividades que se realizaron en este ejercicio. En este caso simplemente vamos a mostrar lo que el doctor Méndez propone. No fue único ese ejercicio de estimación, entendemos, insisto, que hubiera sido mejor algún otro diseño, por supuesto.

Entonces, este es el documento de Nacho y de Hortensia, donde se replica. Ahí tenemos estimadores, base de datos, etc., etc.

En antecedentes también tenemos cómo fue el flujo de la información, cómo trabajamos esa noche. Las tres compañías recibían su información directamente a sus oficinas, las transmitían al IFE, en algunos casos era realmente lo mismo. Las compañías nos daban la información, el IFE validaba.

Los disquetes venían en esos colores para hacer inconfundible el proceso. Había un proceso de filtrado y se repartía en los subequipos que había esa noche del Consejo Técnico. Hay que decir que fue un Consejo

Técnico que estuvo apoyado por gente muy brillante del IFE, del Registro, en cuanto a la cartografía.

Vamos a antecedentes otra vez, por favor Manuel. Algo que es importante, si vemos inconsistencias; en inconsistencias la idea era que la información la veníamos validando cada vez que llegaba una remesa –las remesas llegaban cada 15 minutos– o cada vez que se podía, en algunos casos, identificábamos con esos filtros qué cosas habían pasado: no correspondían los números con la lista nominal, estaba fuera de sitio, etc., etc. Hay una serie de razones que están perfectamente identificadas y explicado el por qué del asunto.

Bueno, la idea es que esto quede como una memoria viva en el mejor de los sentidos, para que la gente que quiera ver, que quiera replicar. Cuando dice Edmundo Berumen que no era una buena muestra y que había que desecharla, bueno, para ser mala muestra no dio mal las estimaciones, no es eso una razón, pero las razones de que sea mala son muy discutibles, es un poco como lo del filtro, no estamos seguros y esa es la importancia de la estadística.

Ahora aquí hay una serie de documentos más o menos informativos, luego vamos a ver carga de remesa. En carga de remesa hemos querido mostrarles a ustedes cómo fue el proceso de la gente, de cartografía, cuando venía validando el asunto. La idea era seguir geográficamente cómo venía todo el proceso. Tienen ustedes aquí a su disposición, o van a tener, todas la remesas, como fueron llegando.

Por ejemplo, yo podía ver si había grandes problemas. En general hubo muy, pero muy pocos problemas, realmente funcionó muy bien el proceso, simplemente es para mostrarles esto, no me quiero detener demasiado.

Bueno, aquí está el reporte de la inconsistencia en ese momento. ¿Qué es lo que se hacía usualmente en este tiempo cuando había una inconsistencia? Se trataba de ver si era grave o no, se corregía, pero entraba en línea inmediatamente al siguiente análisis.

Bueno, ahora uno de los asuntos que era muy importante era ver cómo venía la información a cada uno de los cortes. Entonces, lo que vamos a ver aquí es un elemento que fue fundamental a la hora de evaluar, si podíamos o no salir.

Vamos a ver, por ejemplo, Enrique Alduncin, totalmente al azar, a las 22:02 horas cómo iba su cobertura. Teníamos por casilla; qué tantas casillas de las que se debieron haber capturado en ese momento habían sido ya capturadas y qué tanto representaba eso en términos de lista nominal, de qué gente estábamos hablando.

Entonces, vamos a crear una gráfica. Aquí lo que me va a decir es cómo iba el grado de avance en ese mo-

mento, a las 22:02 horas, respecto de las entidades federativas. La línea roja, en este momento. Era más que esperado el tamaño de muestra que se tenía que cubrir. La verde, en ese momento, era recibida. Esto nos daba ya una idea de cómo venía la muestra a las 22:02 horas.

Aquí está el reporte por entidad.

La idea es que a las 22:02 horas teníamos un reporte por cada una de las empresas geográficamente. Y esto a lo que nos involucraba desde el punto de vista, ya al conjuntar, era una idea de qué tan bien repartida o no estaba la muestra. Para nuestra fortuna, y tenemos que mencionarlo así, era una muy buena repartición a nivel nacional de la muestra. Por eso es que entre otras causas, por supuesto que las estimaciones que se estaban produciendo, tanto al interior del Comité como de las empresas que nos estaban informando cada media hora acerca de sus cálculos, se decidió salir.

Entonces con eso concluiría. Ya no tengo tiempo, pero alguna de las preguntas que esto nos dejaría hacia el futuro van en muchos sentidos: ¿Qué tipo de ejercicios estadísticos podríamos hacer en el futuro? ¿Estos seis años que han pasado respecto a 1994 y que nos han dejado realmente en otro país, podríamos extrapolarlos, sí o no? ¿Tendremos ya dentro de seis años unas elecciones electrónicas, al menos en algunos casos? ¿Qué otros tipos de cuestiones estadísticas? ¿Tendremos o no un centro?

Yo creo que eso es importante por lo que mencionaba Edmundo Berumen: el problema del tiempo y el problema de la muestra se pudieron haber resuelto si hubiera un organismo permanente, un observatorio electoral que no necesariamente tendría que depender tanto del exterior.

Tal vez si tuvieran sus propios recursos, los investigadores podríamos hacer una labor y convencer, tal vez, a nuestros colegas de hacer buenas muestras. ■

PONENCIAS MESA 3

.....

Edmundo Berumen

BERUMEN Y ASOCIADOS PARA IFE

No voy a reseñar en este foro lo que constituye alguno de los distintos ejercicios de investigación electoral. Me voy a conformar, y será un éxito para mí, si simplemente logro dejar planteadas una o dos preguntas, que no espero que sean resueltas aquí, y con eso yo me doy por satisfecho. A lo mejor ustedes quedan frustrados, pero yo me daré por satisfecho.

Para mi beneficio más que para el de ustedes, escribiré algunas fórmulas sencillas, pero a mí siempre me ayuda a regresar a lo básico. Todos recuerdan el estimador de un total basado en una muestra aleatoria simple de tamaño n de una población de tamaño N , que me permito reescribir varias veces:

$$\hat{Y} = N * \bar{y}$$

$$= N \frac{\sum_{i=1}^n y_i}{n}$$

$$= \frac{1}{n} \sum_{i=1}^n \frac{y_i}{\frac{1}{N}}$$

$$= \frac{1}{n} \sum_{i=1}^n \frac{y_i}{p_i}$$

Pero cuando ya llegué a esta manera de reescribirla, al menos yo ya le entiendo un poquito más a la fórmula y digo: ¡Ah, qué fórmula tan inteligente! ¿Qué es lo que está haciendo? Bueno, pues lo que está haciendo es tomar una de las observaciones y la divide por 1 entre el tamaño N de la población, y ¿qué es eso? Ese término de la suma, por sí solo, está estimando lo que busco: el total de la característica de interés. La variable que observé, piensen en votos de la sección electoral i a favor de un partido, lo estoy dividiendo entre 1 sobre el total de los elementos en la población (entre el total de secciones, por ejemplo). En un muestreo aleatorio simple, cada extracción de una observación, de un elemento que voy a observar de la población de interés, es extraído del universo con esta probabilidad, 1 entre el total N de la población.

Entonces, este solo término estima lo que busco, con la simpleza de multiplicar lo observado en una sola unidad por el total de unidades en el universo. Pero como no tengo sólo una unidad medida, tengo n de ellas, estoy sumando cada una de estas estimaciones del mismo total y luego las promedio y, entonces, tengo un promedio de n estimaciones de lo mismo, que luego en modelos de muestreo más complicados generalizamos. Entonces, decimos: para estimar un total, lo único que tengo que hacer es dividir cada observación entre su probabilidad de selección, y luego promediar:

$$\hat{Y}_{ppt} = \frac{1}{n} \sum_{i=1}^n \frac{y_i}{p_i}$$

$$p_i = \frac{C_i}{C}, \text{ medida de tamaño}$$

Llegamos, entonces, a una clase de estimadores que toman las observaciones, dividen entre la probabilidad de selección, las suman y luego las promedian. Un subconjunto de esa clase de estimadores son los estimadores donde la selección de las unidades a medir es con probabilidad proporcional a una medida de tamaño, que llamamos selección con PPT, en el lenguaje del muestreo. Piensen que las medidas de tamaño son ciudadanos del padrón y que estamos seleccionando secciones electorales con probabilidad proporcional a los parámetros del padrón. La varianza muestral de esta clase de estimadores está dada por:

$$V(\hat{Y}_{ppt}) = \frac{1}{n} \sum_{i=1}^n p_i \left(\frac{y_i}{p_i} - Y_{ppt} \right)^2$$

$$\begin{aligned} \hat{Y}_{ppt} &= \frac{1}{n} \sum \frac{y_i}{p_i} \\ &= \frac{C}{n} \sum_{i=1}^n \frac{y_i}{C_i} \end{aligned}$$

y_i = número de ciudadanos que votaron por el partido A

Si se toma la esperanza de ese estimador, pues es simplemente la esperanza de cada uno de sus sumandos que es todos los valores posibles de la población, de la variable que estoy usando, por la probabilidad de haber seleccionado ese elemento, que en el muestreo con PPT es la medida de tamaño. Hace uno el álgebra y llegamos al total poblacional, entonces decimos que es un estimador insesgado. Todos contentos hasta ahí:

$$\begin{aligned} E(\hat{Y}_{ppt}) &= \frac{1}{n} \sum \left(\sum_{i=1}^N \frac{Y_i}{p_i} p_i \right) \\ &= \frac{1}{n} \sum_{i=1}^n \sum_{i=1}^N Y_i \\ &= \frac{1}{n} n Y_A = Y_A \end{aligned}$$

De la misma manera se estima el total de ciudadanos que votaron por cualquier partido, esto se convierte en la variable Xi digamos, y se toma el cociente de estas dos estimaciones para estimar el porcentaje de interés.

Sin embargo, algunos colegas utilizaron otro tipo de estimador para estimar el porcentaje de votos a favor de un partido:

$$\hat{p}_{ppt} = \frac{1}{n} \sum \frac{y_i}{v_i}$$

v_i = número de ciudadanos que votaron

Donde V_i difiere de C_i , de acuerdo a la tasa de participación en cada sección electoral:

Ahora viene la pregunta que quiero dejar como una de las cosas que tenemos pendientes. El estimador usado por algunos colegas es distinto del que detallamos al principio. Aquí el denominador es las personas que

votaron por el partido, que es la clase que decía Raúl Rueda. Cada término de la suma es el porcentaje a nivel de la sección electoral en muestra, es el porcentaje de votos para la Alianza por el Cambio, por el PRI, por el PRD, para cualquiera de estos, en una sección electoral en particular.

¿Qué es lo primero que notamos diferente de la clase que estábamos viendo antes? Que el denominador ya es una variable aleatoria, ya no es un parámetro como en el modelo más sencillo. Antes era una constante, que era la medida de tamaño con la cual seleccionaba la sección electoral.

Esta pequeña diferencia hace que esto ya no sea una constante, el número de ciudadanos en el padrón o lista nominal, sino una variable; es simplemente el número de ciudadanos que en la sección electoral i ejercieron su derecho a votar; fueron, hicieron su cola y votaron. No lo conocemos *a priori* como sí *a priori* conocemos la medida de tamaño, es una variable aleatoria.

¿Qué pasa si yo tomo la esperanza a esa clase de estimadores? Bueno, uno de los viejos trucos que nos enseñan cuando estudiamos estadística, ¿cómo tomo la esperanza de algo bivariado? primero tomo una esperanza condicional, condiciono sobre una de las variables para mantenerla fija, tomo la esperanza sobre la otra variable, y si las cosas me ayudan y se cancelan los términos problemáticos, al tomar la segunda esperanza todo mundo vive feliz.

Entonces, si yo tomo la esperanza de esta clase de estimadores, esto es la esperanza de cada uno de estos términos, la condicional, estoy condicionando en que es conocido el total de ciudadanos que ejercieron su derecho a voto en cualquiera de las secciones electorales, el resultado es el siguiente:

$$E(\hat{p}_{\text{ppc}}) = \frac{1}{n} \sum_{i=1}^N \left(\sum_{j=1}^N \frac{Y_j}{V_j} p_j \right) \\ = \frac{1}{n} \sum_{i=1}^N \sum_{j=1}^N \frac{Y_j}{V_j} \frac{C_i}{C}$$

Donde V_i difiere de C_i de acuerdo a la tasa de participación en cada sección electoral.

Pero una de las inconveniencias del resultado es que los términos molestos no se cancelan.

Queda pendiente ahora tomar sobre estas sumas la esperanza que falta, que es sobre los posibles valores del número de ciudadanos que votan, es una esperanza que todavía está pendiente de tomar. Esto requiere algún tipo de modelo de distribución.

Aquí estamos corriendo riesgos fuertes, cuando estamos utilizando en nuestros sistemas de estimación de ejercicios de conteo rápido estimadores de esta clase, cuya esperanza es ésta: ¿Y luego cómo resolvemos la esperanza pendiente? La variable aquí es 1 sobre el número de ciudadanos que ejercieron su derecho a votar. ¿Cómo se distribuye eso? Podemos empezar a modelarlo, qué les parece que el número de ciudadanos que votaron en la sección electoral i , es una función muy sencilla que depende del tamaño del padrón de la sección electoral, el número de ciudadanos en esa sección por una tasa de participación.

$$v_i = \beta_i C_i$$

Le debemos sumar un término de error a ese modelo.

Si sustituimos este modelo en la expresión me permite cancelar las C_i pero deja el pendiente de: ¿Qué vale tu parámetro de tasa de participación a nivel de cada una de las secciones electorales? Ah, pues sigo simplificando, porque nos encanta simplificar. Bueno, pues vamos a usar otro modelo, vamos a suponer que la tasa de participación es pareja y que en cualquier sección electoral de este país la tasa de participación es la misma:

$$v_i = \beta C_i$$

Entonces, usemos este modelito y, entonces, ya las cosas son muy bonitas porque al sustituir se cancela lo problemático y ya puedo, entonces, decir que esta expresión se reduce, y regreso a un estimador que va a ser simplemente un promedio del total del porcentaje, del total del padrón, multiplicado por la tasa de participación.

Bueno, pero estas simplificaciones y este modelaje –insisto– han ignorado el término de error que estos modelos tienen. Yo anoche le preguntaba a algunos que estaban alrededor de la mesa mientras cenábamos, si una variable se distribuye como una normal, el término de error que me falta poner aquí, ¿cómo se distribuye a la inversa? Algún día lo sabía, ya se me olvidó, llamémosle una normal inversa.

Esto implicaría que agregáramos un modelo adicional sobre el término de error, que nos quedaría pendiente aquí de modelar y de estimar. Ese es uno de los pendientes que tenemos, desde mi punto de vista, para un ejercicio que consideramos que es sencillísimo de una estimación de conteo rápido usando esta clase de estimadores.

Los ejercicios que hicimos en el IFE no usaron esta clase de estimadores. Ya presentó Raúl Rueda la clase de estimadores que usamos en el ejercicio del conteo rápido para el IFE. A diferencia de estos, que son estimadores, que se llaman separados, estimadores de razón separados porque es el promedio de cocientes, donde numerador y denominador son variables, el que para el conteo del IFE, y que puso en una de sus láminas Raúl Rueda, es lo que se llama un estimador combinado, se estima todo, el numerador de manera independiente de la estimación de todo el denominador, por un lado estimo los votos a favor de la Alianza por el Cambio y después, de manera independiente, estimo el total de votos para cualquier partido, más no registrados, más nulos.

Hay varias ventajas que en la literatura son conocidas de los estimadores combinados sobre los estimadores separados cuando el tamaño de muestra a nivel de cada estrato es mayor a 20. No es necesario el recurrir a la demostración, eso está disponible en varios textos.

Sin embargo, a pesar de eso, varios colegas que usaron conteos rápidos usaron el estimador de razón separado, que es menos preciso.

Ahora, si yo logré transmitirles mi inquietud de trabajar sobre este tema, de usar esa categoría de estimadores y resolver lo que no está resuelto de un error adicional que ya proviene del modelo, ya no proviene del muestreo, yo me doy por satisfecho.

Pero quisiera retomar ahora un par de preguntas nada más de la relación de las empresas que hicimos el conteo para el IFE con el Comité Técnico con el que estuvimos trabajando.

Un resultado también harto conocido es que la varianza de un diseño que tiene fijación óptima es menor o igual que la varianza de un diseño que tiene fijación proporcional y es menor o igual que la varianza de un diseño que usa un muestreo aleatorio simple.

Para el tipo de problemas que estamos discutiendo, esta relación se sostiene, es más eficiente una fijación óptima que una proporcional, que un muestreo aleatorio simple.

En el conteo del IFE el diseño, que fue responsabilidad del Comité Técnico, y ya Roy Campos lanzó a la mesa una pregunta que ya no es tanto técnica, ya es de una naturaleza más importante, ¿qué tanta responsabilidad puede asumir la empresa sobre su trabajo si se le quita la responsabilidad de diseñar la muestra? Creo yo que es una pregunta muy interesante que vale la pena discutirse, pero en el terreno técnico la pre-

gunta valedera es por qué voy a usar un sistema de selección de secciones electorales con muestreo aleatorio simple, si sé que son más eficientes los discutidos.

Además, si uso aleatorio simple con proporcional, también es sabido que si se estratifica y se usa sistemático, donde el sistemático en el orden de la lista tiene una estratificación implícita, eso es más eficiente que un muestreo aleatorio simple con afijación proporcional.

Yo creo que en ejercicios tan relevantes como estimar parámetros del resultado de una elección, deben tener especial cuidado de revisar y tomarnos libertades que en otras condiciones no nos las tomamos. Yo creo que el valorar la muestra seleccionada y tener la capacidad de decir: la tiro a la basura porque es mala, es una responsabilidad para este tipo de ejercicios; y selecciono otra y la valoro y digo: sigue siendo mala, la tiro, y así hasta que sale una que es "buena".

Hay varios criterios y varios parámetros para juzgar esto y no creo que es faltarle a la estadística el usar esos criterios para decir el ejercicio de estimación que voy a hacer va a ser con el uso de esta muestra. Eso yo creo que es una de las facetas que en ejercicios futuros valdría la pena discutir. Yo creo que los colegas que no trabajan para clientes como el IFE, en las condiciones de este ejercicio en particular, probablemente varios de ellos lo hicieron, probablemente se tomaron el tiempo de examinar las muestras seleccionadas y desechar varias para quedarse con algunas de las buenas.

En el IFE no hubo esa oportunidad. Se nos dio la muestra asignada. Ya se comentó que fue responsabilidad del Comité Técnico. Yo examiné la muestra particular que a nuestra empresa le tocó y a nivel global se veía muy bien, pero cuando la analizábamos por estado, por distrito electoral, y comparábamos las composiciones urbanas, mixta, rural, de las secciones electorales *versus* la clasificación del padrón, empezaban a surgir discrepancias serias y fuertes que pudieron haber impactado el proceso de estimación, que felizmente no fue el caso; pero fue un felizmente *a posteriori*, cuando se pudo haber prevenido y tener más certeza, un grado mayor de certeza, de que no podrían pasar cosas desagradables.

Creo, entonces, el segundo tema que quiero dejar asentado como tema a discusión es la pertinencia de que en ejercicios tan relevantes, por lo que dijo Raúl Rueda del costo político que tienen, se tome el tiempo para valorar muestras *a priori* y desechar las que nos pueden poner en riesgo el ejercicio de estimación. ■

DEBATE

.....

GUILLERMO BAZ: Pasamos ahora a la serie de preguntas y respuestas. A mí me gustaría hablar de cuatro cosas. El conteo rápido tiene tres ingredientes fundamentales: uno, es el diseño estadístico; el segundo, es la cuestión de logística, que creo que no deben estar separados el diseño de la logística, creo que es parte del asunto. Por otra parte, está el cliente como un elemento muy importante. El cliente está definiendo muchas de las acciones que se están haciendo en los conteos rápidos.

Me gustaría comentar algunos de los aspectos, primero, de los que mencionó Raúl Rueda, en cuanto a diseño. Hay un supuesto que se hace, que es el de intercambiabilidad, que de entrada es un supuesto que no funciona; el orden y, en particular, cómo llegan los resultados al centro de acopio, pues de entrada no podemos suponer intercambiabilidad.

Hay otros aspectos, me parece que en cuanto a diseño al haber definido 300 estratos, eso va un poco en contra del diseño. No siento que está adecuado a las características propias del conteo rápido, son demasiados estratos, se me hace, es mi opinión.

Consideración respecto al tamaño de muestra. Creo que todo mundo estará de acuerdo en que estamos excedidos en tamaño de muestra, que siempre se estuvo excedido en tamaño de muestra. Creo que esto de los 300 estratos es demasiado. Primero, porque es un diseño caro, muy caro, con una dispersión geográfica demasiado grande y creo que es innecesaria; otra cosa es que imposibilita el poder realmente hacer algunas estimaciones un poco más rápido.

En cuanto a Edmundo Berumen, cuando entró a la clase de estimadores, bueno, yo no sé quién está entrando con esa clase de estimadores, pero para mí es una complicación innecesaria; o sea, simplemente se tiene que considerar al no votante como un partido más y se acabó el asunto. Tenemos al PRI, PAN, PRD, otros, y los que no votaron son otro partido, y darle el mismo tratamiento que se le da a cualquiera de los demás partidos.

Entonces, como tema académico podría ser interesante, pero yo creo que no viene al caso meterse a ver cuál es la esperanza de los estimadores y todo eso. Creo que nos complicamos la vida innecesariamente.

En cuanto a lo que comenta Edmundo Berumen, de ejercicios que se hacen, tomo esta muestra y luego no me gusta y a lo mejor luego la vuelvo a cambiar, yo creo que ese conocimiento del no gusto o gusto de la muestra se puede incorporar al diseño; o sea, si el diseño que estás haciendo te permite que te genere muestras que no te gustan, pues mejor incorpora ese conocimiento al diseño y puedes tomar sin miedo cualquier muestra que generes.

De Enrique Alduncin no tengo mayores comentarios. Qué bueno que nos comentó algo de la experiencia, yo creo que de la experiencia se aprende mucho, a veces mucho más cuando las cosas no salen como uno quisiera, que cuando salen perfectamente. Creo que cuando salen muy bien hay una posibilidad de ya no aprender un poco más. O sea, yo creo que el no haber podido presentar los resultados en el momento que se necesitaban, bueno, ese es un error y se debe reconocer como tal, y yo creo que se aprende mucho de esos errores. Y si nos pudieras transmitir algo de la experiencia de lo que te pasó, ya sea en campo, que no te llegaban los resultados, o te falló el campo, no sé qué, es clarísimo que algo falló. Si nos pudiéramos enriquecer de esa experiencia, sería buenísimo.

FRANCISCO ABUNDIS: Yo tengo un comentario más como usuario de esta información, que me parece que en la noche electoral ante la sorpresa o el primer impacto de lo que había sucedido con la victoria de Fox se perdió un poco de vista.

Cuando el IFE da los resultados, los da en rangos, que es una forma muy distinta de darlos en comparación con las otras empresas. Sobre esto de hecho hubo un debate con Warren Mitofsky acerca de cuál era la forma más apropiada de hacerlo, independientemente de si esta fórmula es más o menos confusa, da rangos que dan un número absoluto.

Lo que creo que fue desconcertante para quienes seguíamos esto muy de cerca, es el punto medio de los rangos que se dieron. En el caso de Gallup, que se dio para Fox 44.0 y 48.4, y Berumen dio 41.2 y 45.2—esto es por poner un ejemplo—, el punto medio de Gallup ahí hubiera sido 46.2 y el de Berumen 43.2. El promedio de eso es 44.7. El dato del IFE en su punto medio dando el rango de 39 y 45 era 42, había 2.7 de diferencia.

Yo digo, ahora ya recibí una explicación de, bueno, hay una parte de la muestra que estaba integrada que no estaba entre esas dos empresas, que era la parte de Enrique Alduncin, pero en ese momento era muy desconcertante porque la Presidencia estaba bien definida, pero había otros datos como la Cámara, que si bien era una medida independiente, eso era un buen punto de referencia para ver qué estaba pasando a esta hora, si bien los conteos rápidos ya habían dado mucha información, sobre todo los privados. Esto era, sin duda, el ejercicio con la muestra más grande, según entiendo, en términos de lo que eran conteos.

Esa es una primera observación sobre un debate que yo creo que no terminó de resolverse, primero, sobre si dan rangos o dan un dato preciso y luego sobre el reporte de la información que, sin duda, yo creo que

muy poca gente lo notó y quedó claro quién era ganador y perdedor, pero eso fue simplemente una circunstancia afortunada; pudo no haber sido así.

Quisiera señalar también algunos otros puntos del debate de esos días, y yo creo que días posteriores. Creo que la idea de citar a Warren Mitofsky es útil porque él daba como las críticas de la parte comercial de esto y tal vez no de la más metodológica. Si alguna ventaja tienen sus críticas—creo yo— es que son visiones muy prácticas de cómo reportar esto, de cómo hacer un anuncio electoral. Yo no me meto a la parte de muestra, simplemente a la parte de reporte.

Una de sus críticas era la falta de uso de algoritmos en la estimación. Él decía que a nosotros no nos quedaba claro, pero sabíamos que había algún problema o alguna diferencia y otro era este de la integración, cómo se iban a integrar las muestras, que al parecer es lo que explica esta diferencia.

Entonces, yo creo que parte de la revisión de lo que pasó con los conteos, nos obligaría a regresar a esas críticas de origen de Warren Mitofsky que aparecieron en Milenio, en este artículo, y sería un buen ejercicio, simplemente porque él—creo que— tiene alguna visión y una práctica distinta en esto.

ROY CAMPOS: Aclarando lo que presentó Francisco Abundis, esto es una pregunta al IFE: ¿Cómo integró, cómo realizó los intervalos de confianza política?

Estos son los puntos medios de la empresa uno, puntos medios de la empresa dos y el punto medio de la integración del IFE, los intervalos, límites inferiores. Entonces, no es la unión como nos presentó Raúl Rueda, si esto incluyera a la tercera empresa, tendría que venir ganando Labastida para la tercera empresa 38-36 para que los puntos medios promediaran, bueno no van a ser puntos medios, pero no es una integración ortodoxa de intervalos; este intervalo no da. Si hay una integración de la tercera empresa, bien, no está a la vista, no se dijo que se integró, se dijo que se integraron dos y también no se puede decir que si no se dijo o no se publicó, igual no se publicó lo de Warren Mitofsky; esto es lo que se dijo, se integraron dos, la forma de integrar no está.

En el caso particular, el intervalo de la empresa uno no se intercepta con el intervalo de la empresa dos, éste termina en 36.2 y éste empieza en 36.2. De hecho, el Comité Técnico aquí, una de las dos empresas no iba a incluir su estimación en el intervalo, no hay intersección en el intervalo para Labastida.

Entonces, aquí hay una duda de cómo se integró este intervalo. Cuando Raúl Rueda lo dice, dice: se integró como unión. Bueno, ni como unión.

La otra pregunta que es prácticamente el cuestionamiento principal de lo de Warren Mitofsky o lo mío también es: si una autoridad electoral debe contratar. Lo que veíamos nosotros fuera del IFE porque, bueno, este es un cliente político, era: viene una elección cerrada, lo veíamos, había consenso, salvo en algunos encuestadores y la autoridad electoral va a tener ese día, bueno, en esa jornada, dos resultados.

Si los medios nos equivocamos, nos equivocamos, somos malos; si salimos y nos equivocamos, al día siguiente el IFE dice sale esto y qué malos encuestadores son los medios. Pero si el IFE dice un ganador a las 23:00 horas y al día siguiente sale otro, no se equivocó, hizo fraude; o sea, el IFE iba a tener dos resultados.

Entonces la pregunta es: si en una elección competitiva, como se veía, ¿debía la autoridad electoral hacer un conteo rápido? Esa era la pregunta.

La siguiente pregunta es: la autoridad electoral contrata, nosotros como gremio ahora nos preguntamos ¿cuando nos contrate una autoridad electoral nos impone el diseño, o sea, nos impone el diseño y nosotros vamos, levantamos y presentamos ese resultado y lo firmamos con nuestro nombre? Porque esto es lo que contestó Berumen en Milenio, y no se vale decir que ahora es de otra manera, ya apareció.

El Consejo Técnico le cuestionó la metodología y terminaron haciendo otra. Eso fue lo que pasó. Son las preguntas que dejo.

JAVIER ALAGÓN: En realidad las dos preguntas fueron planteadas con uno u otro matices. Y en relación al tema que tocó Edmundo Berumen, me gustaría abundar un poco en mi presentación del conteo rápido en la tarde.

Sólo quisiera agregar que, desde luego, ya habíamos considerado el aspecto de si no se tienen las mismas tasas de votación pueden ocurrir algunas desviaciones; sin embargo, creo que las ventajas de estratificar y de utilizar el muestreo con probabilidad proporcional al tamaño (PPT) sobre un muestreo aleatorio simple, creo que son bastante claras y por eso optamos algunas empresas, muy seguramente otras también, por este tipo de diseños, aun sabiendo que ocurría eso.

Nada más tratando de comentar sobre esta última pregunta que hizo Roy Campos, sí me gustaría añadir que sí hay empresas que nos cuestionamos si debíamos participar en un ejercicio como hizo el IFE. Yo tuve una invitación por parte de COPARMEX para hacer el conteo, y de manera simultánea hubo un acercamiento por parte del IFE, y parte de la decisión –parte, no toda– de no participar con el IFE fue, precisamente, el riesgo de no haber tomado la muestra.

Entonces, desde luego que también de repente sí se toman en cuenta esos factores.

ANA CRISTINA COVARRUBIAS: De mi parte nada más una aclaración hacia Rubén Hernández, que tiene que ver con esto que se ha estado diciendo de que compartamos bases de datos. No sé si están enterados, pero si no es el caso, por eso lo quiero ratificar, que desde el inicio de los trabajos de TV Azteca, que tuve la responsabilidad de realizar, hubo un compromiso de TV Azteca de entregar la base de datos, a fin de que fuese confirmable totalmente y, bueno, sometida a, si era necesario, un tipo de auditoría y más que nada, sobre todo, que contribuyera a ser analizada y a este intercambio de información que se está pidiendo, y creo que es el mismo caso de quienes trabajaron para CIRT.

Entonces, la base de datos del conteo y las encuestas de salida de TV Azteca fueron entregadas al IFE dos días después de la elección, así que ya obran en poder de ustedes, y lo mismo pasa con Bimsa.

IVÁN CASTRO: En primer lugar, yo no estaría de acuerdo con un comentario que hizo Rubén Hernández con respecto a que el diseño de muestra en un conteo rápido puede representar un problema tan difícil como la selección de filtros o ponderaciones en la cuestión de una encuesta preelectoral.

Yo creo que aquí tenemos el panorama mucho más fácil desde el punto de vista estadístico y sí debe de existir por ahí el diseño óptimo, al final de cuentas todos trabajamos con la misma población, con la misma información, entonces debe de estar por ahí el diseño óptimo, sin muchas dificultades para encontrarlo.

Con respecto a este tema, tal vez sí deberíamos de enfatizar un poco más el por qué –y ya se ha mencionado– se definieron tantos estratos, ¿por qué se definió una muestra proporcional para cada estrato? Si generalmente las muestras cuando se define el número de elementos en cada estrato óptimo funciona mucho mejor, ¿por qué el muestreo aleatorio simple sobre las secciones? Bueno, pero son cuestiones que se pudieran discutir.

Lo que sí se me hace un poco grave es que la misma autoridad académica, en este caso, que tuvo el IFE, transmitiera la idea de que la comparación de intervalos era la forma directa para conocer quién era el ganador. Sabemos que los porcentajes que se estiman son porcentajes que están correlacionados, que suman 100. Entonces, no es el traspase de intervalos, sino es una prueba sobre la diferencia, eso es importante.

El enfoque bayesiano me pareció excelente, es algo en lo que se podría trabajar más, incluso porque como a diferencia de muchos otros casos, yo creo que aquí se tienen bastantes elementos para definir distribuciones *a priori* correctas o muy aproximadas a lo que va a suceder al final.

El comentario que hace el doctor Edmundo Berumen con respeto al muestreo proporcional al tamaño (PPT) y al estimador, yo creo que es muy valioso. En efecto, *a posteriori*, ya cuando tenemos los datos de la información y si hacemos simulaciones, sí existe un sesgo en hacer un PPT sobre la lista nominal y después considerar exclusivamente al total de votantes.

Acá existe un gran beneficio, que es que existe una gran correlación entre la lista nominal y el total de votantes en la sección, entonces al final las empresas que decidieron realizar un PPT sobre lista nominal, bueno, tenían un seudo PPT sobre el total de votantes y por eso no existieron grandes problemas. Pero el comentario es muy valioso, y como lo dice Guillermo, el problema tiene una solución simple y rápida, considerar la abstención como un candidato más.

También el comentario que hace el doctor Edmundo Berumen sobre la estratificación, bueno, que podemos abundar un poco más.

Otra cosa. Yo creo que sí podemos definir pruebas de dispersión que no sean exclusivamente gráficas. De hecho, en Bimsa estuvimos manejando una prueba de dispersión que se basaba en grupos de dispersión de la misma muestra, por el mismo método de selección que era sistemático, a través de una prueba de bondad de ajuste. Entonces, también son cuestiones que se pueden analizar desde un punto de vista de prueba estadística.

MIGUEL BASÁÑEZ: Entiendo que el conteo rápido y encuesta de salida de ustedes quedó en el lugar 11 de 16, si no estoy mal, no sé si eso lo platicaste tú con Warren. Sin embargo, de esto me parece que hay una cuestión importante: quién debe de contratar las encuestas de salida y los conteos rápidos, los medios o la autoridad electoral, el IFE.

Lo estoy pensando más que nada para los estados, si los IFE's estatales van a continuar con esta práctica y si habría los medios locales que pudieran hacerlo, y creo que aquí hay una reflexión que sería de importancia para nosotros.

Sólo una aclaración: Mitofsky no es el creador de las encuestas de salida en los Estados Unidos, es Bob Louis, el que era director de Encuesta de Los Angeles Times, y después se incorporó Warren y después siguió, pero nada más como un punto informativo.

MANUEL ROMÁN: Muy constructivamente. Primero, una observación de Enrique Alduncin, que nadie se quejó del uso de las encuestas como una salida de mercadotecnia política. Yo estuve atento a la radio, a la televisión y en recintos públicos más limitados y en otros abiertos.

Hubo muchos de nosotros que no sólo nos quejamos, sino que quisimos dar un mensaje propedéutico al televidente o al radioescucha de ese uso, de ese manejo, varios están aquí.

Al comentario de que la encuesta del vecino no me dejó satisfecho, diríamos que una de las cosas que hemos ganado es que anteriormente varios de nosotros decíamos quién está a cargo y, entonces, la confianza se da; se sabe quién es y, entonces, se respeta la manera como se trabaja, y se aprecia. Muchas de las encuestas de mis vecinos sí me dejan satisfechos, inclusive desde luego las del Comité Técnico, las del IFE, las de Edmundo Berumen y las de Roy Campos.

En ese sentido, creo que sí tenemos esos elementos. Por eso, para ir más allá de lo que es la personalidad, nos gusta ir hacia los detalles metodológicos, y voy a mencionar los que corresponden como inquietudes personales más acerca de lo que se acaba de presentar.

Primero, yo estoy muy satisfecho, y felicito al Comité Técnico del IFE de la experiencia de la vez pasada, de tres réplicas independientes de mil 500 secciones, que sí eran independientes, se construye un modelo que va calibrando la información.

Sabemos muy bien que modelos que son o buscan la optimalidad, desde el punto de vista muestral, en la práctica, cuando se divorcia del trabajo de campo, no sólo no son óptimos, sino llegan a la suboptimalidad, y ese es parte del problema, muchas veces de no respetar lo que es de diseño y no operar con la logística apropiada.

La pregunta directa sería hacia el grupo técnico y hacia los encargados de hacer este estudio, si en algún momento estuvieron monitoreando también la información a través de otra técnica, tipo replicaciones, o sea, con grupos aleatorios. Si fue el caso, si de manera parcial utilizaron pruebas de banda de ajuste como acababan de mencionar.

FRANCISCO SÁNCHEZ VILLARREAL: Bueno, una de las preguntas iniciales que se hicieron en esta mesa y que de alguna manera complementa una de las proposiciones y cuestionamientos que dejaron los compañeros del CEO en las presentaciones de la Mesa 2, es sobre la forma de reportar la vitrina metodológica. El caso más particularizado en este punto es, precisamente, la forma en que se reportan los valores de esos porcentajes de votos que se están estimando.

Si se hace la exclusión de los que no contestan, si se da únicamente una estimación puntual, si esta estimación puntual debe de acercarse, cerrarse al punto porcentual más cercano o puede, en un momento dado, considerar décimas de punto porcentual, e incluso hasta centésimas, con la falsa pretensión de que al tener más decimales se está siendo más exacto. A veces se interpreta incorrectamente esta situación.

Finalmente, en cuanto a la expresión o la complementariedad –diría yo– de plantear esos intervalos de confianza, esto es, dar un estimación puntual, creo que no se divorcian la una de la otra, sino que plenamente se complementan.

Desde luego, también, la crítica que yo considero muy válida y que la hemos tomado en consideración anteriormente, antes ya la había expuesto Iván Castro en el Colegio Nacional de Actuarios, en cuanto a ese falso concepto metodológico de proponer los intervalos de confianza manejando las proporciones como si fuesen independientes, y hay por ahí algunas publicaciones que se han filtrado tratando de difundir esta situación, precisamente, de manera totalmente incorrecta.

Creo que esos son puntos fundamentales. No tenemos que irnos hasta los refinamientos que nos propone Raúl Rueda, elegantes y satisfactorios, en el momento que concluyen, con un estimador de razón combinado, pero son detalles elementales que –creo yo– son los que debemos de ir tomando en consideración para definir, no reglas, sobre qué muestreos se deben utilizar y si son muchos o pocos estratos, tratando de aplicar un poco la crítica de Guillermo Baz.

Recordemos que la estratificación nos garantiza, precisamente, dada una variable que esté amplia o deseablemente correlacionada con los resultados, el evitar muestras que aleatoriamente hubieran quedado mal distribuidas. Una mala muestra puede concentrarse de manera errónea y la estratificación finalmente nos va a llevar a que esa situación se corrija.

Finalmente, si analizamos lo que se pretende hacer con las estratificaciones, pues, entonces es deseable, y metodológicamente comprobado, ya que entre más estrato se tenga, si es que no se necesita eventualmente y las muestras que se están manejando son lo suficientemente grandes, para que los grados de libertad que se pierden en las estimaciones no nos causen mayor problema hacerlo. Es –creo yo– bastante deseable, y más en muestras como las que institucionalmente requiere el IFE, donde lo mencionó claramente Raúl Rueda, se requiere oportunidad pero también una extraordinaria precisión, porque cualquier desvío ligeramente sustancial se va a interpretar, aun cuando el orden de los candidatos o partidos se respete, se interpreta como

una cierta preferencia de la institución electoral y como consecuencia de parcialidad hacia un determinado candidato o partido.

Estos son los aspectos que ciertamente pueden normar, pero también, por otro lado, sí, la libertad metodológica es la que seguramente va a redundar en una mayor creatividad. Es totalmente absurdo, desde mi punto de vista, coartar la libertad metodológica de las empresas y que sea, precisamente, la experiencia y foros de esta naturaleza, en los cuales se discutan los detalles metodológicos a partir de los cuales todos iremos aprendiendo en el sentido de no solamente aprender la estadística; no hablamos únicamente de su aprendizaje, sino de ese aprender, esto es, de su aprender que solamente nos proporcionan los datos y el manejo continuo de grandes cantidades de datos.

El Comité Técnico manifiesta, pues, la dificultad que tuvieron para poder emitir resultados; los resultados en los cuales tenían responsabilidad, pero mucho más responsabilidad tenían las empresas, y más con el mal sabor de saber que esos resultados no eran totalmente suyos, una vez que la metodología les había sido impuesta de una manera u otra.

Imaginemos también algo que depende del cliente, en este caso es la autoridad electoral, pensemos en el tipo de situaciones que se plantean cuando se trabaja para medios en los cuales los resultados no se quieren tres o cuatro horas después de haber cerrado las casillas, sino media hora antes de que éstas cierren. Son circunstancias totalmente distintas, pero finalmente tenemos que dar respuesta a la demanda de los diferentes clientes.

No podemos tampoco, por la ortodoxia técnica y científica, condicionar a lo que es el deseo del cliente, sino llegar a un justo medio, adecuado, que a ambos proporcione confianza.

Calificar los resultados cuando estos han sido soportados metodológicamente de manera correcta, reduce también la posibilidad de malas prácticas. Se hizo algún comentario de las desviaciones que pudo tener Nielsen, a pesar de toda la ortodoxia técnica que utilizó; bueno, creo que esa es la tranquilidad de ellos, tuvieron una mala muestra y se acabó; pero si metodológicamente estuvo correcta, esto es una situación de excepción y no una situación generalizable.

Precisamente, esa situación es lo que permite que podamos calificar que un trabajo es honesto, confiable y que no necesariamente de manera sistemática se van a presentar desviaciones. Creo que hay muchos más grados de libertad o licencias poéticas que se utilizan en una práctica cotidiana de muestras, que por pura casualidad le pegan de todas maneras al bueno, y eso

no es una garantía total de que se esté trabajando correctamente.

EDMUNDO BERUMEN: No voy a contestar lo que han preguntado, eso me lo reservo para la réplica, e invito a mis colegas que hagan lo mismo. Yo lo que quiero es simplemente agregar un par de observaciones.

Uno, la diferencia sustancial entre la manera de exhibir los resultados del IFE por las razones expuestas ya por Raúl Rueda, deberían de tener la misma estructura de resultados que después conoceríamos vía PREP, de la manera de exhibir resultados del resto de los conteos en general, con algunas excepciones.

El IFE dio la distribución de los votos para cada alianza o partido más los no registrados, más los nulos, a diferencia de la inmensa mayoría de estimaciones de fuera que quitaban a los no registrados y a los nulos.

Entonces, una primera confusión de ese día eran las estimaciones, la distribución de los conteos contratados por el IFE –no del IFE– *versus* los conteos rápidos externos. Entonces, creo que ese es un tema que vale la pena comentar para, a futuro, ver si es recomendable una u otra acción al respecto.

Lo segundo, es que no deja de ser apreciado como alguna aberración estadística, aunque se pueden construir poblaciones donde se cumpla el que a mayor muestra crezca la varianza.

Si el ejercicio de consolidación es: voy a consolidar muestras, y luego ese ejercicio de consolidación me lleva a intervalos más grandes, me está creciendo la varianza. Existe, y podemos construir ejemplos de texto para presentar en clases poblaciones donde sucede que, con mayor muestra aumenta la varianza, pero en situaciones de la inmensa mayoría de ejercicios, y en particular de este ejercicio, es una aberración estadística.

ROLANDO OCAMPO: Refiriéndome al anuncio de la publicación del sábado 1° de julio en el diario Milenio acerca de Warren Mitofsky. El problema –creo yo– no era parte tanto del aspecto técnico, sino cómo vino a enrarecer el ambiente también en esos momentos. Estábamos ante un escenario, yo trabajé para la Alianza por el Cambio, todo mundo lo sabe, donde había un ambiente muy adverso hacia el candidato de la Alianza, donde los medios de comunicación se fueron cerrando y había muchas críticas, muchas notas negativas hacia el candidato, y el decir que las muestras estaban sesgadas a lo urbano; entonces, favorecen a Acción Nacional, entonces, ¿el IFE es Foxista?

Ese era fundamentalmente el problema a la gente, no a los técnicos como somos nosotros, me incluyo, sino al medio nacional, a todos los lectores, como

estaban enrareciendo el ambiente, y justo Milenio la publica el 1° de julio. Eso yo creo que es fundamental.

Una cuestión más, yo sí quisiera que aclararas, Rubén Hernández, si fue un acto de mala fe del IFE o de José Woldenberg el no haber anotado a Enrique Alduncin en los dos conteos rápidos, porque creo que el aspecto que comenta Enrique Alduncin de que lo tienen marginado, a lo mejor también llegó hasta allá con José Woldenberg o, efectivamente, se está automarginando.

En ese sentido, yo quisiera, por eso lo quiero comentar porque habló, y pides una disculpa, Enrique Alduncin, y das una definición de lo que es el sesgo matemático, pero luego en cuanto aparece la nota de Gauss, que por cierto aparecen otras empresas como GEA, como Edmundo Berumen, y luego apareces tú también, me vuelves a decir que sí estoy sesgado. Entonces, ya no entiendo cuál fue la definición de tu sesgo.

ROY CAMPOS: No hablé de las cosas buenas, literalmente, para los objetivos que tenía el IFE, y que presentó Raúl Rueda, cumplió totalmente sus objetivos; o sea, le dio certidumbre al proceso. El anuncio acabó la jornada electoral y en ese momento cumplió el objetivo para el que planteó el IFE.

También, la parte del diseño que presentó Raúl Rueda, para no ser cursi, simplemente voy a decir que está muy buena, la veo segura.

Sobre la pregunta de Miguel, sí pidió disculpas. No nada más pidió disculpas, lo que dijo ya lo había dicho, se lo dijo a José Woldenberg antes, se lo dijo a Juan Molinar antes, y sí hubo una disculpa ese día temprano, tanto al IFE como a Televisa, que era el cliente, como a mí, que era su contratante; o sea, sí hubo disculpas al respecto.

Sobre de que fuimos 11 de 16, no sé a qué te refieres de esa clasificación, si fue la 11 que nos registramos, la 11 en precisión, cuando es lo que menos se califica. En mi presentación voy hablar de los objetivos de un conteo, certidumbre, oportunidad, eficiencia, y la última es precisión. Si estás clasificando en precisión, no fuimos la 11, estábamos en un nivel de audiencia (*rating*), bajo ese criterio, pero no son los criterios para nosotros que se deben de calificar.

Las precisiones, estoy de acuerdo con Rolando Ocampo, el momento oportuno (*timing*) de la entrevista fue muy malo; o sea, eso es lo criticable. Si ahora apareciera, fuera una discusión metodológica y no tuviera ningún impacto.

Doctor Edmundo Berumen, si dice que lo que se publica, que considera como un partido, además los no registrados y los nulos, con más razón esos intervalos –como mencionaba–. Bueno, creo, y si no los

puntos medios de una de las empresas de tres partidos sumaban 102.5, faltaba todavía incluir el otro; o sea, iba a sumar con 106, 107, con más razón están mal los intervalos o los puntos medios de los intervalos.

ENRIQUE ALDUNCIN: Primero, me vuelvo a disculpar, Ana Cristina Covarrubias, en la última reunión, en el IFE, dio incluso por nombre las empresas que estaban falseando la información y, desde luego, a los medios, pero no fue una cosa de gremio, sino fue una cosa personal de Ana Cristina Covarrubias, y fue una gotita de agua que se perdió en la marea de información de la guerra de encuestas que estaba dentro de la guerra política. A eso me referí y que, en general, nos dejó dañados.

Respecto a lo del conteo rápido, pues, sí teníamos nosotros como empresas libertad y, desde luego, pudo haber sido mejor el ejercicio, aunque pienso que fue espléndido. Nos pudieron haber citado antes, pero también tengo entendido que la decisión de hacer un conteo rápido fue una decisión del Consejo de última hora y no hubo tiempo para planearlo e invitar a las empresas.

Entonces, nosotros aceptábamos o rechazábamos, ya dijo Javier Alagón, que él rechazó, el GEA también rechazó.

Esta es una gráfica que me presta Rubén Hernández de los resultados, aquí estamos presentando los resultados. Aquí la situación es la estadística clave, es la diferencia entre el primero y el segundo lugar. Esa es la estadística clave, incluso está señalado aquí con negritas las 22:00 horas; entonces, a las 22:00 horas yo tenía casi el 60 por ciento de la muestra, Berumen tenía como el 45 y Gallup iba bastante adelante, con un 72 por ciento de la muestra.

Entonces, no fue que no hubiera cumplido con el registro de campo, que era la asignación fundamental que se nos hacía, ya que nosotros no participamos en el diseño del experimento.

Aquí lo que es importante es que, a pesar de que estábamos usando el método *Clusters* que ya se mencionaba y con los ponderadores que tratan de corregir el efecto de que, primero, llegan los resultados urbanos y luego los rurales, lo que pasó en 1994, por lo que se cayó el sistema.

Aquí vemos que nosotros teníamos que cuidar mucho como empresa porque también el IFE fue muy ingenioso al contratar a las empresas, no era el IFE el que estaba dando los resultados eran las empresas. La prueba es de que a pesar de que tenía el IFE un indicador más exacto que el nuestro, puesto que estaba sumando las tres muestras, dio a las empresas por delante.

Entonces, aquí vemos en mi caso a las 22:00 horas, que fue el corte que se usó, yo tenía ocho puntos de diferencia entre el primer lugar y el segundo, cuando una hora antes tenía 12 puntos de diferencia.

Entonces, nosotros teníamos que ver cuándo se estabilizaba la convergencia. Yo ya sabía que había un ganador, sin embargo, este margen se iba a seguir cerrando, la prueba es que cerró en 6.5, yo tenía aquí ocho.

Como se ha mencionado, el conteo rápido es una prueba de oportunidad y de precisión, de las dos, no es como el chiste del español de qué prefieres: oportunidad o precisión, tienen que ser las dos. Y en mi caso yo veía que todavía no estaba perfectamente cerrada la brecha; sí veía que iba ganando claramente, ya sabíamos que había ganado Fox, pero por cuánto.

Entonces, yo tuve mi libertad de decir, como empresa: no considero que deba dar los resultados, quiero ver que la convergencia vaya y, de hecho, se cerró todavía un poco más.

Aquí vemos que en el caso de Gallup tenía una brecha de cinco y, en el caso de Edmundo Berumen, tenía una brecha todavía de nueve puntos de diferencia cuando se hizo el anuncio.

Entonces, no teníamos, a pesar de que los intervalos no se cruzan –como se comentaba–, para mí la estadística fundamental, más que los intervalos, era la diferencia en las medias, y aquí yo no veía todavía una convergencia clara y podía hacer, siendo una cosa tan delicada como la elección presidencial, de dar un resultado; o sea, aquí el error de tipo dos, el de comisión, era más grave que el de omisión.

Por ese motivo yo dije: tengo que esperar a tener más muestra, a pesar de que tenía una muestra ya bastante razonable.

Yo creo que lo que sí estuvo un poco mal y fue confuso, es el que no se mencionara que dentro de los resultados que yo sí había cumplido, porque la impresión que se dio a mucha gente es que no había cumplido con el requisito fundamental que era levantar el trabajo de campo, ya que nuestra participación en el diseño fue cero, simplemente, aceptar o rechazar lo que ya habían hecho.

Yo, por mi parte, hice ejercicios para ver esta situación y acepté la muestra. Me pareció espléndida, me pareció un diseño muy ingenioso, muy bien armado, y la prueba está en los resultados, aunque aquí lo que más importa es la metodología, más que la realidad.

RAÚL RUEDA: Antes que nada, había comenzado mi exposición diciendo que no sabía si dar las gracias o no, ahora sí quiero dar las gracias porque este tipo de discusiones me encantan.

Son más preguntas que el día 2 de julio, pero voy a tratar de contestarlas agrupándolas. Básicamente, las puedo agrupar en tres: una es con respecto al diseño; dos, desde el punto de vista académico y, otra, desde el punto de vista no académico. Las voy a llamar así: una con respecto al diseño, otra a la forma como se consolidó la información y la otra con respecto a la imposición del diseño.

Debo decir que el IFE, el Consejo General, más bien el Presidente del Consejo General, como atribución propia que el Consejo General le dio, decidió contratar a tres empresas, a las tres empresas contratadas; es decir, nosotros no tuvimos nada que ver.

El mismo Consejo General decidió que las empresas presentaran sus metodologías y decidió que el Consejo Técnico decidiera qué hacía con ellas. El diseño que nosotros pensamos fue para el Consejo Técnico, si lo quieren ver así. Nosotros lo que teníamos que hacer era: combinar la información, que no quedó muy claro, pero se vio en esa lámina cuando me dijeron que faltaba un minuto y puse esa lámina ahí, que no salen las cuentas porque los promedios no dan; bueno, es obvio, porque no es un estimador lineal lo que nosotros estamos usando; lo que hacemos es una combinación lineal de distribuciones de probabilidad.

El valor esperado de esas cosas no es una combinación lineal simple, que es el promedio de los estimadores puntuales o de las medias de las distribuciones, si no van ponderados por las probabilidades de cada uno de los estratos, son problemas condicionales. Entonces, no salen por eso.

Que los puntos medios de los intervalos de probabilidad, bueno, nosotros lo que dimos no fueron intervalos de probabilidad, quiere decir que no necesariamente son simétricos.

Lo de la unión, bueno, a veces hablo libremente, en sucio, hablo como pienso. Al decir unión, era un criterio más o menos minimax, más bien en ese sentido, lo más peor que podía pasar, eso fue lo que hicimos.

Ah, pero antes de eso, ¿por qué se salió con intervalos? Eso fue decisión de las empresas, así querían dar los resultados al IFE. Yo les mostré en la lámina que lo que nosotros hicimos fue encontrar una distribución multivariada, que es la distribución final, sobre esa encontramos la distribución de la diferencia condicionada a los votos válidos, que era la gráfica que les decía, donde evaluábamos la probabilidad de cometer errores. Esa es una distribución sobre un solo parámetro, pero el parámetro de interés, que es la diferencia entre los votos. En efecto, hay un efecto de dependencia entre los intervalos y por eso construimos la distribución final, precisamente por eso.

Una confusión, como menciona Enrique Alduncin, con respecto a su participación, bueno, el Comité Técnico no es que se quiera zafar, pero nosotros trabajamos en un lado y José Woldenberg estaba en otro lado, hubo un problema de comunicación, el correo perdió ahí algo. Eran las 10:15 horas cuando estábamos consolidando y nos estaban arrancando las hojas. De hecho, algunos de nosotros no nos enteramos que Enrique Alduncin no salía hasta que José Woldenberg lo dijo; o sea que ni idea teníamos, estábamos separados en varios grupos porque había que hacer demasiadas cosas.

Ya hablé del diseño. Los tamaños de muestra, y eso era porque eran tres réplicas, en realidad lo que se estaba haciendo, que fue muy costoso.

Que demasiados estratos, eso se platicó también con el Consejo General y ellos querían que fuera nacional; es decir, que estuviera distribuida, les da un poco de temor que no salieran secciones de Sonora, por ejemplo. Entonces lo que había que hacer era distrital y, mientras más, si tenían el dinero para hacerlo, pues mejor.

El que tengamos más información o menos información no hace el análisis de ésta ni más rápido ni más lento, digo, a las velocidades que tienen las computadoras hoy en día analizar mil o dos mil datos es una diferencia de milésimas de segundo, que yo no alcanzo a distinguir.

Finalmente, con respecto a la imposición. Me llama muchísimo la atención la palabra de imposición porque se discutió el diseño con las empresas antes de que firmaran contrato, y las empresas firmaron contrato. Entonces, ¿es una imposición? No sé si el contrato venía con una condicional o algo así.

EDMUNDO BERUMEN: El comentario de varias de las personas que hicieron algunas preguntas ilustra lo malo que soy para explicarme, y por eso yo me daría por satisfecho si, simplemente, dejara bien planteadas las preguntas que quise dejar planteadas. Voy a intentar explicarme un poco mejor.

Uno, las ventajas de seleccionar muestras con probabilidad proporcional, una medida de tamaño, son muy conocidas desde los años cuarenta, y eso no está siendo cuestionado.

No se están cuestionando las bondades, las ventajas y virtudes de seleccionar las nuestras con una medida proporcional al tamaño. Lo que fue cuestionado es la clase de estimadores que se utilizaron, con un diseño donde la muestra fue seleccionada con PPT, y hablé y traté de recordar la clase de estimadores de razón separados *versus* la clase de estimadores de razón combina-

dos; la clase de estimadores de razón combinados para este tipo de problemas son mejores y no tienen ningún problema en utilizar la probabilidad de selección PPT, tal cual, la medida de tamaño tal cual, sin usar en el denominador los votantes, porque se estima por separado el numerador del denominador y luego se hace el cociente. No hay dificultad técnica alguna.

El problema es cuando usan un estimador de razón separado, donde están tomando el cociente de la sección uno, el cociente de la sección dos, donde el denominador de esos cocientes es con las personas que fueron a votar y en nada resuelve el problema el que le agreguen los nulos, los no registrados y los abstencionistas. Tan no ayuda que subestimarían absolutamente todo, porque lo que están haciendo es incrementando el denominador, están regresando a la lista nominal y están estimando los porcentajes de votación en función de la lista nominal, una aberración total que en nada resuelve, todo lo contrario, tiren a la basura a quien vaya a hacer eso.

El IFE cuestionó –como lo dije– a través del Comité Técnico los procedimientos, metodologías que presentaban las empresas y, en particular, la que presentamos nosotros. Sí, si lo dije –lo repito– sí lo cuestionó y sí se hizo un ajuste. Nunca se dijo, nunca lo dije, no lo digo ahora, que el ajuste que el Comité Técnico sugirió fuese un diseño que no permite inferencias válidas.

Desde mi punto de vista entonces y hoy, y lo repetí, es menos eficiente, punto. Es válido, tiene buena precisión. Pensamos ayer y pensamos hoy que nuestro procedimiento de selección de muestra hubiese proporcionado un estimador aún más preciso, que no es el criterio.

ENRIQUE ALDUNCIN: La precisión de un diseño, como el que sugeríamos, y era exclusivamente para la selección de la muestra, no se refería a ningún otro aspecto. El Comité Técnico, de hecho, aprobó yo diría que el 99 por ciento de la propuesta metodológica, excepción hecha única y exclusivamente en la manera de seleccionar las muestras al interior de los distritos electorales.

Sí, si fue cuestionado; sí, si fue modificado. Y para mí, entonces y hoy, resultaba igual un método aceptable, simplemente menos preciso en su diseño, que pensamos y creemos todavía hoy, que el que sugerimos.

El tamaño de muestra y lo costoso de la investigación, 850 secciones electorales por cada réplica, pero quisiera poner sobre la mesa un elemento adicional, las estimaciones que ya han puesto varias veces y no quiero volver a poner, que tienen cortes horarios.

Vieron por ahí, en el último renglón de cada una de las empresas, el porcentaje de muestra que había.

El porcentaje de muestras que se sugirió no era porque la estimación se iba a dar con esas 850 secciones electorales, era porque la mejor estimación que se tenía del volumen de información con que se contaría a la hora que se nos exigía entregar un resultado, que se nos pedía tener un resultado a las 21:30 horas, para que fuese posteriormente consolidado, para que después se enviase a las 22:00 horas a redactar el informe del Presidente que iba a salir a las 23:00 horas. Entonces, no son datos de las 23:00 horas, son datos de las 21:30 horas, en términos del ejercicio.

Ese tamaño de muestra estaba calculado en función a las 21:30 horas, qué porcentaje de la muestra espero ya me habrá sido reportado, y con base en eso se especulaba que tendríamos del orden de un 60 por ciento de la muestra a la hora que se iba a demandar, que por ahí vieron andaba en cincuenta y tantos por ciento en una empresa, nosotros la más baja. Un paréntesis ahí.

¿Por qué el volumen de información de Berumen, no recuerdo, andaba en treinta y tantos o cuarenta y cinco por ciento, por qué era mucho más baja que mis colegas? Bueno, pasa lo que pensamos no va a pasar, y pasó.

En nuestro caso, por razones hoy todavía no entendibles, se desconectó el enlace de Edmundo Berumen con su empresa, y el volverlo a conectar tuvo una traba que hoy día no se sabe el origen; alguien estaba colgado en ese enlace, que no era ni Edmundo Berumen ni era personal del IFE, ¿Quién? Hoy todavía no lo sabemos.

Ese alguien que estaba colgado, al no cortar la comunicación, impedía que restauráramos el enlace. Por eso es que en la estadística de la empresa ven que el volumen es mucho menor a los de mis colegas que, qué bueno que a ellos no les sucedió eso, pero en nuestro caso eso pasó.

Incorporar al diseño el cambiar las muestras con base en una evaluación *a priori* que sugiere Guillermo Baz, pues sí, es justo lo que estaba solicitando, es factible hacerlo y no tiene mayor dificultad.

RUBÉN HERNÁNDEZ: Seguramente ya hay muchas cosas que fueron respondidas, algunas ya se me olvidaron, me dejaron al último y ya se me olvidaron, tal vez empezar por las más próximas aquí en mis notas.

En efecto, Rolando Ocampo, yo no podría decir que hubo una mala voluntad en contra de Enrique Alduncin. Realmente, y debo transcribirlo, porque es la verdad, fue un tanto caótico esos últimos tres minutos, donde el papel tenía que salir hacia la oficina

de José Woldenberg, fueron caóticos. Y, en efecto, nosotros, el equipo que estábamos trabajando, todo el tiempo estuvimos conscientes de que la aportación de esta compañía estaba presente en nuestros cálculos. Eso para nosotros era transparente.

El cómo ese mensaje no llegó allá, eso es parte -digamos- de nuestro patrón, en este caso, que me sirve esto como puente. Yo no hablo por el IFE, no soy más que un empleado, que se acabó mi contrato en ese momento, nadie me obligó a hacer cosas que no quería, estuve con mucho gusto.

Creo que, como mis colegas, y no podía dejar de hablar de Mitofsky, él dice que el IFE no es su negocio hacer conteo rápido, tiene toda la razón. En efecto, el IFE responde a otras necesidades, ya las dijo Raúl Rueda, tener una población que está expectante, informada, aún con esos elementos; yo, que sepa, la estadística no fue hecha única y exclusivamente pensando en una cuestión.

Lo que no entiendo es ¿por qué dudar que el IFE puede hacer eso? Evidentemente hay riesgos, todos vivimos de eso, bendita variabilidad, lo dije hace rato. Pero, justamente, la idea era cómo este Comité Técnico iba a combinar y, ciertamente -y por ahí respondo a algunos-, pues, no seamos tan inocentes; pues, de dónde les sale a algunos la idea de que nada más teníamos que promediar, eso lo programamos en Excel y se resuelve el problema.

No era ese nuestro papel, por supuesto. Y si hemos entendido lo que significa un intervalo de confianza, pues, no tiene por qué ser el medio, si no saldríamos por el medio, hasta la teoría clásica nos dice que puede ser en cualquier lugar.

Vamos entendiendo, entonces, eso. Por supuesto, claro que no era sumarle y dividíle, no, claro que no. ¿Qué era? Bueno, pues, de aquí aprovecho yo la ocasión, seguramente haremos un taller donde vamos a abordarlo, a recalcular, calcular y a mostrarles todas las experiencias que hicimos.

Hubo -insisto- no solamente lo que se les mostró acá el ejercicio del doctor Méndez respecto a la estimación, hubo la propuesta también de Raúl Rueda; hubo una propuesta muy interesante de O'Reilly, donde él calculaba justamente la probabilidad de que la tendencia se invirtiera; es decir, nos fuimos por la diferencia también.

Hubo otro elemento importantísimo en términos de que, cuando la muestra iba llegando, qué pasaba con esos elementos que no llegaban, no llegaba Sonora, no llegaba Chiapas, ¿qué íbamos hacer? Tienes toda la razón, Iván Rivadeneyra. Por supuesto que hay maneras de medir eso, el problema es que no nos daba

tiempo. El tiempo fue muy tremendo en nuestra contra. Evidentemente, están esas estadísticas y nos hubiera encantado hacer alguna cosa para pronosticar. No tuvimos ni siquiera en mente, antes de este ejercicio, todo lo que pudimos haber hecho.

Yo creo que queda la invitación para que si esto se vuelve a repetir, quien quiera que lo haga, pueda tener estos elementos, ese famoso disco compacto (CD), que por ahí van a tener ustedes alguna vez, tiene esa idea, de que esta experiencia no se quede con los que estuvieron ahí en la oficina del IFE calculando esto, eso es muy importante. Y, por supuesto, son bienvenidas esta y todas las propuestas que ustedes puedan tener.

Francisco, por supuesto que no estuvieron obligadas las compañías. Parece que estaban encerrados con riesgo de tambos y, por supuesto, en todo momento estuvieron libres de decir no, podían salir.

Por supuesto que no fue ese el asunto, no lo veamos así. Yo creo que eso no se vale. Yo creo que todas las compañías que trabajan acá tienen experiencia con clientes difíciles; este cliente sólo les pidió que aceptaran su muestra, que podían probar. Yo creo que tus clientes a veces te piden que es para esta tarde y no le pones: me coaccionó, porque, por favor, todos tenemos mucha experiencia en este asunto.

Lo de las bases, realmente insisto, yo respondo por Rubén Hernández pero, por supuesto, esto veré si el IFE nos la puede pasar, o si no, te pedimos que nos las envíe, por favor.

Yo creo que son muy válidos todos estos elementos de análisis, son muy válidos. Yo creo que los debemos centrar sobre todo en el hecho de lo que se mencionó; Roy Campos, yo creo que lo sintetiza muy bien; bueno, fuimos contratados, estas compañías y nosotros para dar certeza, certidumbre en este ejercicio y se logró.

Yo creo que una elección así no la habíamos ni siquiera soñado, ya no digo hace seis años, hace dos meses no nos la imaginábamos de esa importancia. No digo que este ejercicio vino a darle ese panorama hermoso, por supuesto que no, pero contribuyó. Yo creo que todas las compañías, y nos estamos aquí peleando con un diseño, que si éste sí es, que éste no es, y éste sí le pegó. Ayer nos peleábamos porque no le pegaban y ahora le estamos también echando porque sí le pegaron.

Quisiera yo, pues, moderación. Esto estuvo muy bien, aún para los que no le atinaron a los resultados, ustedes saben que es difícil, los que le atinamos pues también.

Finalmente, por supuesto que puede ser una aberración algunos cálculos, algunas propuestas, lo que pasa es que partimos de ciertos supuestos. A nosotros se

nos podría ocurrir, pero sólo como una hipótesis de trabajo, que a lo mejor las varianzas venían calculadas muy pequeñas, digo, cabe dentro del aspecto metodológico. ¿Por qué suponer que tus varianzas eran las buenas y que las que hicimos acá fue una aberración? Tal vez la aberración hubiera sido traerse unas varianzas, por decir, cualquier cosa, a lo mejor estaban perfectas, no lo dudo.

El hecho es que aquí se hicieron otros cálculos, con otros supuestos y esas varianzas no se hizo, insisto, ni la media, ni solamente ponderar tal cual.

Yo creo que tenemos que celebrar esto, por supuesto esto es jubilariorio, deberíamos estar muy alegres. Por mi parte, creo que este ejercicio fue muy aleccionador. El IFE creo que debe participar muy seriamente en esta línea y la opinión de todos ustedes es muy válida, reunámonos y no dejemos pasar esta oportunidad para participar en serio, y capaz que saquemos el método IFE.

ANA CRISTINA COVARRUBIAS: Una sola aclaración. Ya no hablo, obviamente, a nombre del IFE, pero sí hablo en función de lo que me consta que ocurrió con el IFE y dos cosas que quisiera dejar muy claras.

Una es en relación a por qué el IFE hizo los conteos rápidos, en la AMAI tuvimos un evento un mes antes,

en donde una de las ponencias fue presentada por el maestro Jose Woldenberg, en ese momento José Woldenberg dejó muy claramente establecido que no era el papel del IFE hacer conteos rápidos, que era cuestión de la industria en general y que, bienvenidos todos los conteos que se pudieran organizar.

Lo que ocurrió entre ese evento, que fue el veintitantos de mayo y un mes después, es que hubo una demanda importante de los partidos políticos a través de los consejeros electorales que le pidieron al IFE que volviera a hacer conteos para que hubiera información oportuna en una contienda en la que estábamos viendo que, por lo menos esa era la expectativa, iba a ser sumamente cerrada y que el problema de un conflicto postelectoral estaba claramente planteado.

Entonces, a mí me consta, por lo menos estas declaraciones y, finalmente, esta acción del IFE, que es una decisión de Consejo.

Segundo punto, en cuanto a que si el caso de Enrique fue una cuestión manejada de manera poco informada o de una manera mal intencionada, creo que Enrique nos lo acaba de decir con toda claridad; y parafraseo a Enrique: usé mi libertad y tomé la decisión de que se dijera que con lo que yo traía no me sentía totalmente seguro de cómo iban las tendencias, y eso fue lo que dijo José Woldenberg. ■