

Jornadas de análisis de las metodologías relacionadas con el proceso electoral de 1994

Diseño de muestras para auditorías y conteos rápidos y errores de muestreo asociados

Edmundo Berumen

El trabajo reseña la importancia del diseño y selección de muestras para ejercicios de auditorías técnicas al padrón electoral, así como aquellos cuyo objetivo es la estimación del resultado de la elección en base a los llamados “conteos rápidos”, subrayando los aspectos más relevantes que se deben cuidar. La discusión del trabajo se ilustra con las decisiones tomadas en los diseños utilizados por el IFE para ambos ejercicios en el proceso electoral de 1994.

La estimación de los errores de muestreo asociados a los diseños utilizados es un tema que con demasiada frecuencia se *ignora*, se *olvida*, se calcula de forma equivocada, o se malinterpreta.

El trabajo aborda su pertinencia tanto en la

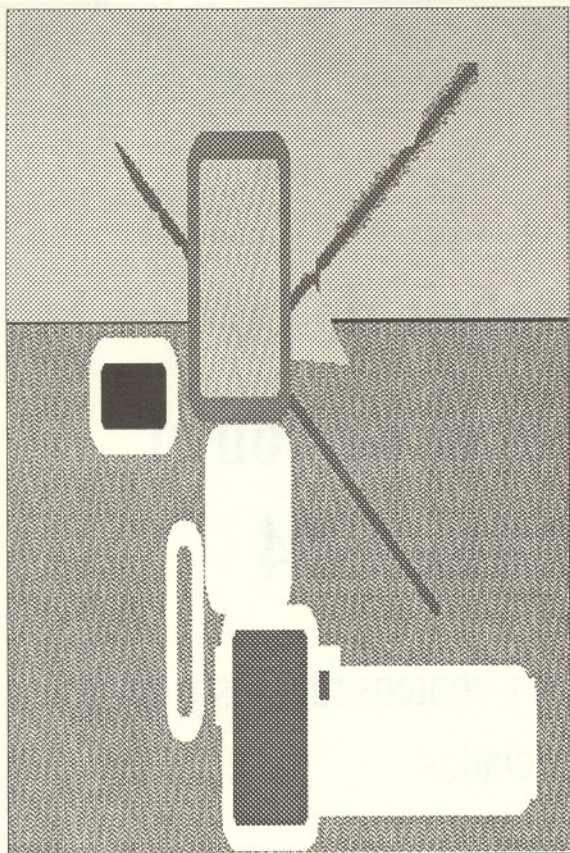
etapa de diseño como en la de estimación de resultados. La discusión del trabajo se ilustra con algunos de los resultados de los ejercicios llevados a cabo para el IFE tanto en la auditoría técnica del padrón electoral como en el conteo rápido.

Por último, el trabajo discute la importancia de varios parámetros metodológicos del diseño de muestra, cuya “portabilidad” es importante para diseños futuros que aborden los mismos temas, así como para diseños en otras áreas cuyas variables prioritarias tengan un comportamiento similar en cuanto a los efectos de “conglomerar” la selección de los elementos a observar.

I. Variables y parámetros a estimar

1. Las variables a estimar en las auditorías técnicas del padrón electoral generalmente se re-

Director General de Berumen y Asociados S.C.



“Ventanas”, (1996), Luis César

fieren a la calidad de los datos consignados en los registros de ciudadanos empadronados, y son fundamentalmente porcentajes para distintas categorías de “inconsistencias” (y por complemento de “consistencias”).

Las inconsistencias que usualmente se estiman son las siguientes:

a. El porcentaje de domicilios no localizados de ciudadanos en el padrón, donde la búsqueda se concentra en la sección electoral referenciada, y en una vecindad cercana del domicilio consignado en el padrón.

b. El porcentaje de ciudadanos no localizados en el domicilio registrado en el padrón, domicilio que sí fue localizado en campo.

c. El porcentaje de ciudadanos no localizados que corresponden a los domicilios que tampoco fueron localizados, del tipo descrito en el punto a.

d. El porcentaje de ciudadanos localizados en un domicilio distinto al del padrón, pero dentro de la misma sección electoral de referencia, en una vecindad cercana al domicilio consignado en el padrón.

e. El porcentaje de ciudadanos que cambiaron de domicilio, según los nuevos ocupantes del mismo o vecinos.

f. El porcentaje de ciudadanos registrados en el padrón, y que ya fallecieron, según los ocupantes del domicilio consignado en el padrón.

g. El porcentaje de ciudadanos que declaran no tener la credencial (no haberla recogido), cuando que el padrón tiene registrado ya haberla entregado.

h. El porcentaje de ciudadanos en el padrón electoral para los cuales existe un “expediente” documental (solicitud de inscripción, de cambio de domicilio, de corrección de datos, de reposición de credencial, o recibo de credencial).

i. El porcentaje de ciudadanos en el padrón electoral para los cuales existen diferencias entre los datos consignados en su credencial y/o expediente documental, con aquellos que el ciudadano da al ser entrevistado durante la auditoría (e.g. sexo, edad, nombres).

2. Se subraya que las auditorías técnicas con trabajo de campo detectan “inconsistencias” que surgen de discrepancias entre lo declarado por los ciudadanos seleccionados al responder el cuestionario de la auditoría, con aquello que está consignado en el registro del padrón electoral. El que los datos discrepen puede tener su origen en distintas situaciones:

a. El ciudadano (o quien dio sus datos) se equivocó al dar el dato al momento de registrarse en el padrón.

b. El ciudadano (o quien dio sus datos) se equivocó al contestar la pregunta del cuestionario de auditoría.

Cuadro 1

Tamaños de muestra bajo un diseño de
muestreo aleatorio simple sin reemplazo

d	Confianza 95% (a = .05, k = 1.96)				
	P=.05	P=.10	P=.25	P=.40	P=.50
.005	7,299	13,830	28,812	36,880	38,416
.010	1,825	3,458	7,203	9,220	9,604
.020	456	864	1,801	2,305	2,401
.030	203	384	800	1,024	1,067
.040	114	216	450	576	600
.050	73	138	288	369	384
.005	5,173	9,801	20,419	26,136	27,225
.010	1,293	2,450	5,105	6,534	6,806
.020	323	613	1,276	1,634	1,702
.030	144	272	567	726	756
.040	81	153	319	408	425
.050	52	98	204	261	272

c. En el proceso de incorporación del ciudadano al padrón se cometió algún error de transcripción del dato documental al dato en archivo magnético.

d. Desactualizaciones no reportadas por el ciudadano.

e. Se dieron una o más de las situaciones anteriores.

Las verificaciones que realizan las auditorías generalmente no pueden detectar cual de las situaciones anteriores es el origen de la discrepancia que genera la inconsistencia.

3. Por su importancia específica y el interés partidista, es frecuente que los diseños de muestra permitan el estimar las inconsistencias para cada una de las entidades federativas.

Y para no empañar el análisis comparativo a nivel entidad federativa, así como cumplir con el

nivel de precisión y confianza exigidos a nivel estatal, el tamaño de muestra y método de selección en las distintas entidades federativas deberá ser igual.

4. En el caso de los llamados “conteos rápidos”, el interés fundamental es estimar el porcentaje de votos válidos emitidos a favor de los candidatos de los principales partidos en la contienda electoral. Cuando la elección es la presidencial, el interés se centra en el resultado nacional.

II. Tamaño de muestra

5. Para definir el tamaño de muestra, se parte de las especificaciones de precisión y confianza que deben tener las estimaciones de los parámetros prioritarios a estimar. Es común que el “cliente” demande orientación para definir la precisión y

Cuadro 2
Valores de $deff_x$
según número de ciudadanos seleccionados y coeficiente roh_x

roh_x	$deff_x$			
	b: Número de ciudadanos seleccionados por sección electoral			
	20	30	40	50
0.00	1	1	1	1
0.02	1.38	1.58	1.78	1.98
0.04	1.76	2.16	2.56	2.96
0.06	2.14	2.74	2.34	2.94
0.08	2.52	3.32	4.12	4.92
0.10	2.90	3.90	4.90	5.90

Cuadro 3
Valores de $deff_x$
según votos válidos y coeficiente roh_x

roh_x	$deff_x$				
	b: Promedio de votos válidos por sección electoral				
	100	300	450	600	700
0.00	1	1	1	1	1
0.05	5.45	15.95	23.95	30.95	35.95
0.10	10.90	30.90	45.90	60.90	70.90
0.15	15.85	45.85	68.35	90.85	105.85
0.20	20.90	60.80	90.80	120.80	140.80
0.25	25.75	75.75	113.25	150.75	175.75

confianza que requiere. La responsabilidad de quien va a realizar la auditoría es dársela de la manera más explícita y clara posible.

6. Formalmente, para determinar un tamaño de muestra inicial se parte del supuesto de un muestreo aleatorio simple, y de las exigencias de precisión y confianza requeridas, expresados en términos generales por:

(1)

$$PR [| p - P | \leq d] > 1 - \alpha$$

Donde :

P = porcentaje real de discrepancia-inconsistencia entre los datos proporcionados por el ciudadano (o por el personal de campo en el caso de domicilios y ciudadanos no encontrados) y los datos correspondientes contenidos en el registro del padrón, para los distintos atributos de interés; o el porcentaje de votos válidos emitidos a favor de un partido en el caso del conteo rápido.

p = porcentaje que estima la investigación.

d = máxima diferencia aceptable entre el valor real, **P**, y su estimación, **p**, prove-

niente de la investigación.

$1 - \alpha$ = confianza requerida en la precisión, o sea el cumplimiento de la declaración probabilística enunciada en la ecuación (1).

7. La fórmula (1) y las exigencias de precisión y confianza conducen a la siguiente expresión para la estimación inicial del tamaño de muestra:

(2)

$$d^2 \geq k^2 * o_p^2$$

Donde **k** se determina en función de la confianza requerida, o sea, según el valor especificado para $1-\alpha$.

8. De (2) y el supuesto de un muestreo aleatorio simple se deduce que se requiere entonces un tamaño de muestra inicial de al menos:

(3)

$$N_0 = \{k^2 * P * (1-P)\} / \{d^2\}$$

9. El tamaño de muestra inicial que se obtiene vía la fórmula (3), se determina bajo el supuesto de

Cuadro 4

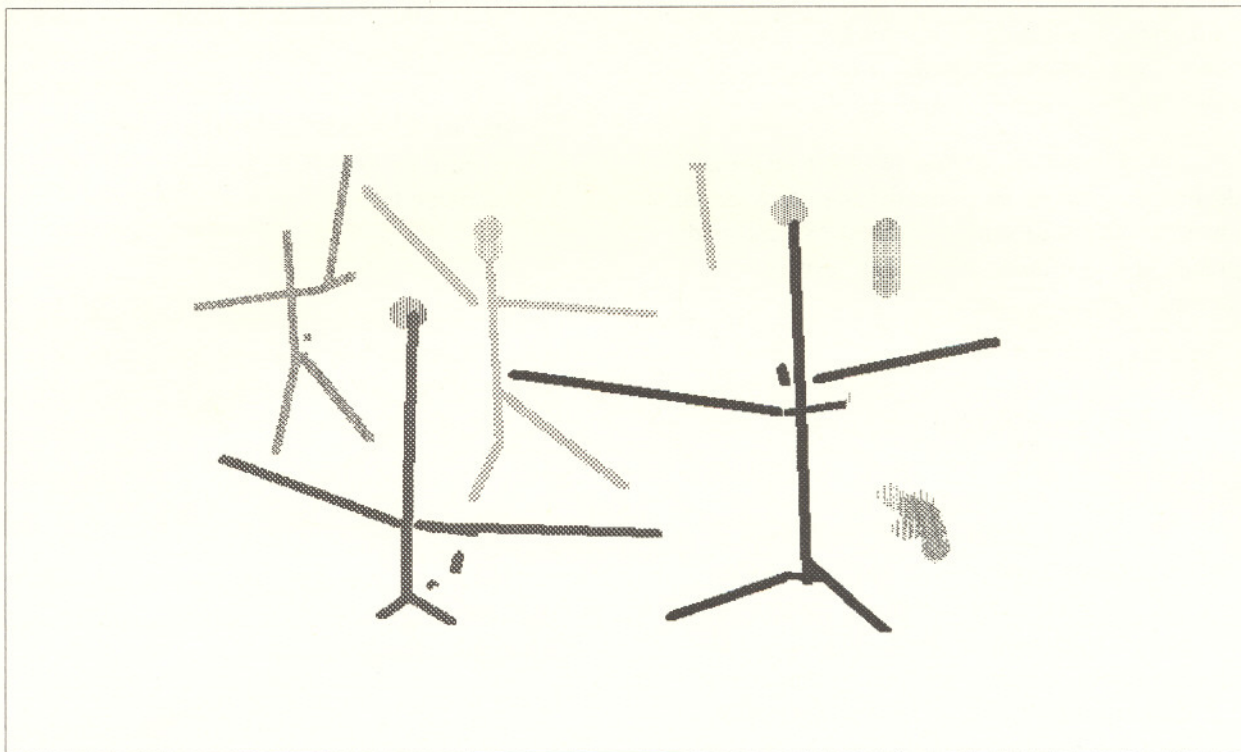
*Equivalencias en secciones electorales
asumiendo un promedio de
396 votos válidos*

Precisión “d”	Nivel de confianza del 95 %		
	P = 0.25	P = 0.40	P = 0.50
.005	73	94	98
.010	19	24	25
.015	9	11	11
.020	5	6	7

Cuadro 5

*Equivalencias en secciones electorales
tomando en cuenta un $d_{eff}=46$
(secciones con igual probabilidad)*

Precisión “d”	Nivel de confianza del 95%		
	P = 0.30	P = 0.40	P = 0.50
.005	3,358	4,324	4,508
.010	874	1,104	1,150
.015	414	506	506
.020	230	276	322



“Los fantásticos”, (1996), Luis César

un muestreo aleatorio simple sin reemplazo. De hecho, tanto para las auditorías técnicas como para los conteos rápidos, lo común es que se utilicen esquemas de muestreo estratificados, con **dos** etapas de selección para el caso de auditorías técnicas (primero secciones electorales y después ciudadanos dentro de ellas), y con **una** sola etapa en el caso de conteos rápidos (secciones electorales, tomando todas las casillas que contengan), al interior de cada estrato.

Lo anterior implica la necesidad de modificar el tamaño de muestra inicial para reflejar la utilización de un diseño distinto al aleatorio simple.

10. Para tener una idea de los órdenes de magnitud involucrados aún con el esquema de un muestreo aleatorio simple, se presenta el resultado de aplicar la fórmula (3) para distintos valores de $1 - \alpha$, d y P :

11. Valores comunes para d en ejercicios de auditorías técnicas varían entre **.005 - .020**, ya que los porcentajes a estimar son generalmente pequeños (P es del orden de **.025 - .150**); en tanto que para conteos rápidos es común que los valores de d oscilen entre **.015 - .040**, ya que el porcentaje de

votación a favor de los principales partidos varía entre **.25 - .60**. El nivel de confianza exigido por lo general es del 95 % (o sea $\alpha = 0.05$). Cuando se hacen estimaciones nacionales y para dominios de interés (e.g. entidades federativas), es común aceptar una confianza menor para estos, generalmente del 90 % (o sea $\alpha = 0.10$).

12. Es tarea del muestrista guiar al “cliente” en la selección del tamaño de muestra inicial apropiado para la precisión y confianza que se requieren.

13. Sin embargo, para lograr la precisión y confianza que se definan, **no** bastan los tamaños de muestra indicados en el Cuadro 1, pues están basados en un esquema de muestreo aleatorio simple.

Es necesario por tanto ajustarlo tomando en cuenta la eficiencia que se pierde al recurrir a un muestreo distinto (estratificado, y conglomerado por el uso de secciones electorales), reflejada en el coeficiente del efecto de diseño conocido como “ $deff_x$ ” donde el subíndice “ x ” subraya el hecho de que el coeficiente depende de la variable/parámetro que se esté estimando.

14. Se sabe que una aproximación del $deff_x$ está dada por:

(4)

$$deff_x = 1 + roh_x * (b - 1)$$

donde:

roh_x es el coeficiente de correlación intraclase ("rate of homogeneity") dentro de la sección electoral (unidad de muestreo o conglomerado de selección), que mide para cada x cuán homogéneo es ésta dentro de la sección electoral (parámetros de inconsistencia en el caso de las auditorías técnicas, porcentajes de votación en el caso de conteos rápidos), y

b es el promedio de unidades observadas dentro del conglomerado (sección electoral), ciudadanos en ambos casos.

15. La relación anterior permite especular sobre los valores posibles de $deff_x$ para distintos valores de roh_x y distintos valores de b . En los Cuadros 2-3 se presentan los valores resultantes de aplicar la fórmula (4) para distintos valores de roh_x y b .

16. En el Cuadro 2 se hace el ejercicio para el caso de auditorías técnicas al padrón electoral, donde de cada sección electoral se seleccionan entre 20-50 ciudadanos para la investigación, o sea las opciones de b oscilan entre estos valores, pero una vez decidido es fija para toda sección electoral en muestra. Los valores de roh_x utilizados son bajos dado el tipo de variables/parámetros investigados en las auditorías técnicas.

17. En el Cuadro 3 se hace el ejercicio para el caso de conteos rápidos, donde de cada sección electoral se toman los datos de votos válidos consignados en las actas de escrutinio y cómputo de todas sus casillas. En este caso b no es fija ni es una decisión del muestrista, depende de la participación ciudadana en la elección. Para el análisis se toman distintos valores para el promedio de votos válidos por sección electoral. Los valores de roh_x analizados son mayores por la naturaleza de los parámetros a investigar.

18. El muestrista debe ahora usar toda su experiencia para seleccionar el valor adecuado de $deff_x$ de los Cuadros 2-3 para corregir el tamaño de muestra inicial seleccionado del Cuadro 1. A la luz del tamaño final debe revisar las implicaciones presupuestales y corroborar con el cliente la suficiencia de recursos o modificar los requerimientos de precisión y confianza.

19. La tarea no termina ahí. El muestrista debe prever ahora la **no-respuesta** esperada en el trabajo de campo, y compensar el tamaño de muestra para que la tome en cuenta.

En las secciones siguientes se ilustran los conceptos anteriores.

III. Ilustración en el caso de auditorías técnicas

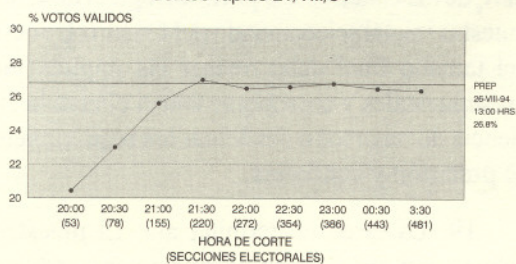
20. En el caso de **Berumen y Asociados S.C.** de acuerdo a experiencias en distintas verificaciones del padrón electoral para distintas entidades federativas, donde se utilizó un diseño de muestra estratificado con dos etapas de selección, se estimó, entre otros, los siguientes "efectos de diseño" ($deff_x$):

	" $deff_x$ "
* para domicilios no localizados dentro de la sección electoral	2.47-2.85
* para domicilios localizados, pero no los ciudadanos	1.42-1.50
* para domicilios que no son viviendas, están deshabitados, o informan que el ciudadano se cambió	1.68-1.69

21. En base a los parámetros anteriores, tomando en cuenta la importancia de la estimación de "domicilios no localizados", para la "auditoría externa" al padrón electoral realizada para el

GRAFICA 1

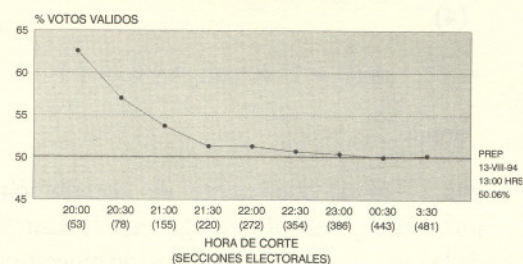
CORTES DE ESTIMACION-PAN
Conteo rápido 21/VIII/94



BERUMEN Y ASOCIADOS S.C.

GRAFICA 2

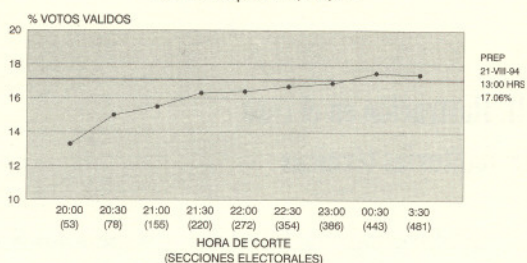
CORTES DE ESTIMACION-PRI
Conteo rápido 21/VIII/94



BERUMEN Y ASOCIADOS, S.C.

GRAFICA 3

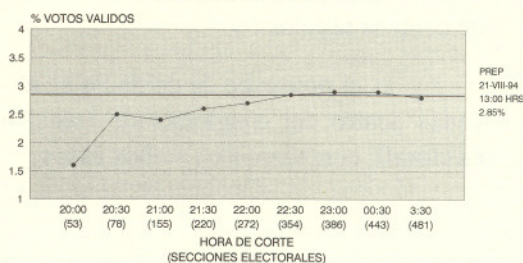
CORTES DE ESTIMACION-PRD
Conteo rápido 21/VIII/94



BERUMEN Y ASOCIADOS, S.C.

GRAFICA 4

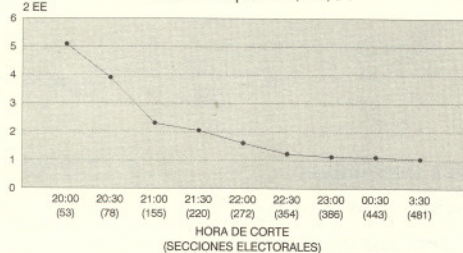
CORTES DE ESTIMACION-PT
Conteo rápido 21/VIII/94



BERUMEN Y ASOCIADOS, S.C.

GRAFICA 5

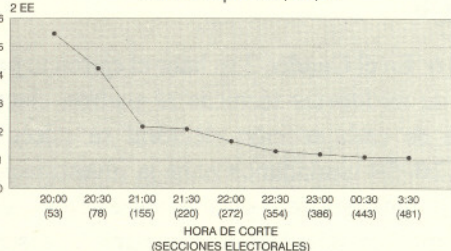
ERROR ESTANDAR DE LAS ESTIMACIONES PARA EL PAN
Conteo rápido 21/VIII/94



BERUMEN Y ASOCIADOS, S.C.

GRAFICA 6

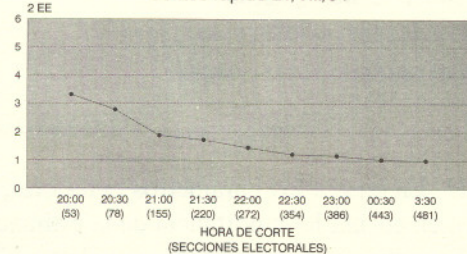
ERROR ESTANDAR DE LAS ESTIMACIONES PARA EL PRI
Conteo rápido 21/VIII/94



BERUMEN Y ASOCIADOS, S.C.

GRAFICA 7

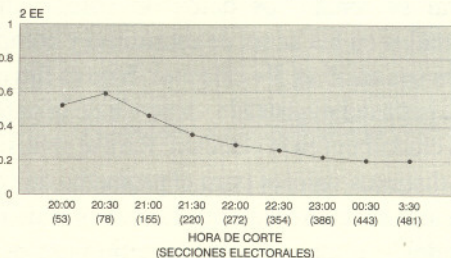
ERROR ESTANDAR DE LAS ESTIMACIONES PARA EL PRD
Conteo rápido 21/VIII/94



BERUMEN Y ASOCIADOS, S.C.

GRAFICA 8

ERROR ESTANDAR DE LAS ESTIMACIONES PARA EL PT
Conteo rápido 21/VIII/94



BERUMEN Y ASOCIADOS, S.C.

IFE, a nivel de diseño se decidió utilizar un *deff* general de 2.9 para modificar un tamaño de muestra inicial de $n_0 = 475$ ciudadanos, lo que llevó a:

$$n'_0 = 475 * 2.9 \approx 1,380$$

22. De acuerdo a experiencias anteriores, la **no-respuesta** esperada en este tipo de investigaciones (ciudadanos que se niegan a conceder la entrevista), es del orden de 2–5 %. Para compensarla, el tamaño de muestra para cada entidad federativa se ajustó tomando el límite superior a:

$$n''_0 = 1,380 \div 0.95 \approx 1,450$$

23. Por último, se deseaba que la información de campo fuera directa del ciudadano seleccionado, pues se intuía que era de mayor calidad. Auditorías previas han logrado información tanto directa del ciudadano como indirecta por parte del cónyuge, padre/madre, hermano(a), otros parientes cercanos. El porcentaje de información lograda en forma directa oscila entre el 67–80%. Finalmente se decidió un valor intermedio, el 72.5%, como el porcentaje esperado de respuesta directa por el ciudadano, lo que implicó un tamaño de muestra final por entidad federativa de:

$$n = 1,450 \div 0.725 = 2,000$$

24. El diseño de muestra al interior de cada entidad federativa seleccionó 50 secciones electorales con probabilidad proporcional al tamaño de su padrón, de una lista ordenada según tamaño y tipo (urbana, rural, mixta), tomando una muestra sistemática de 40 ciudadanos de cada sección en muestra. El diseño produce una muestra **autoponderada** en cada entidad federativa. Por otro lado, el tamaño de muestra global (nacional) asegura una precisión de al menos 0.4 % con una confianza del 95%, para los resultados a nivel nacional de parámetros de inconsistencia cuyos valores estén alrededor del 5 %. (De hecho la precisión lograda fue del 0.2 % para estimaciones de este orden de magnitud, del 5.0 %).

IV. Ilustración en el caso de conteos rápidos

25. De acuerdo a datos del IFE, el tamaño final del padrón electoral arrojó un total de 45,729,057 ciudadanos en las listas nominales de 63,589 secciones electorales. Esto daba un promedio de 720 ciudadanos por sección electoral. Encuestas recientes previas a la elección sobre intenciones de voto arrojaban cifras de participación (ciudadanos que piensan ejercer su derecho a votar) que oscilaban entre 75–87%.

26. A efectos del cálculo de tamaño de muestra **Berumen y Asociados S.C.** asumió un valor conservador para este parámetro y utilizó el 55% como el factor que aplicado al promedio de la lista nominal daría el promedio de votos válidos por sección (más adelante se verá como afecta una mayor participación).

27. Una simple división de los datos del Cuadro 1 entre el promedio de ciudadanos en lista nominal por sección electoral multiplicado por el factor ($720 * .55 = 396$), daba la equivalencia en número de secciones electorales. Estas se presentan, “redondeadas” siempre “hacia arriba”, en el Cuadro 4, para una confianza del 95% y distintos valores para la precisión *d*.

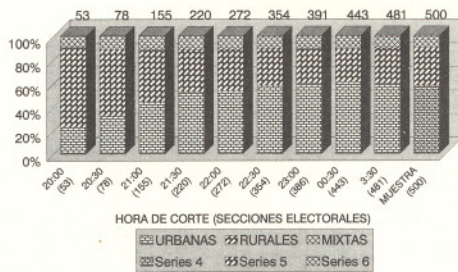
28. Por otro lado, la experiencia de **Berumen y Asociados S.C.** en elecciones recientes había estimado valores de roh_x que variaban según el partido entre 0.056 y 0.090, al interior de secciones electorales.

En esta etapa del diseño se tomó 450 como un número promedio razonable para los votos válidos por sección electoral, lo cual permitía corregir los datos del Cuadro 4 con un *deff* general del orden de 46.

29. Con base en el análisis anterior, **Berumen y Asociados S.C.** propuso una muestra final entre 460 y 550 secciones electorales **por réplica** (el IFE contrataría a tres empresas para realizar tres ejercicios independientes de conteo rápido). Esto permitía asignar una réplica a cada una de las tres empresas participantes, asegurando una precisión mínima del 1.5% con una confianza del 95% con tan sólo una réplica. Dos de ellas aseguraban una

GRAFICA 9

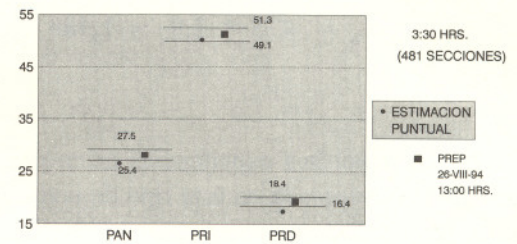
COMPOSICION URBANO-RURAL
Conteo rápido 21/VIII/94



BERUMEN Y ASOCIADOS, S.C.

GRAFICA 10

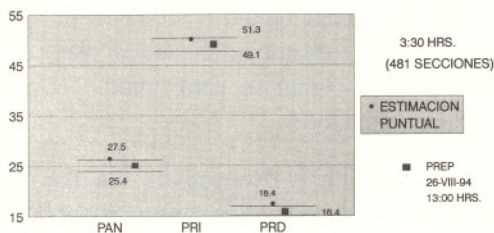
INTERVALOS DE CONFIANZA PAN-PRI-PRD
Conteo rápido 21/VIII/94
PARA LAS TRES PRIMERAS FUERZAS ELECTORALES



BERUMEN Y ASOCIADOS, S.C.

GRAFICA 11

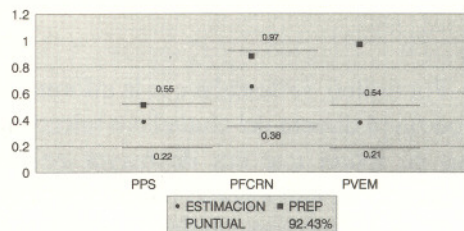
INTERVALOS DE CONFIANZA PAN-PRI-PRD
Conteo rápido 21/VIII/94
PARA LAS TRES PRIMERAS FUERZAS ELECTORALES



BERUMEN Y ASOCIADOS, S.C.

GRAFICA 12

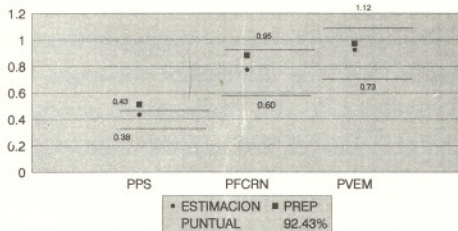
INTERVALOS DE CONFIANZA PPS-PFCRN-PVEM
Corte de las 20:00 Hrs. (53 Secciones)



BERUMEN Y ASOCIADOS, S.C.

GRAFICA 13

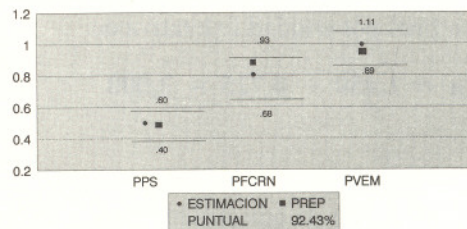
INTERVALOS DE CONFIANZA PPS-PFCRN-PVEM
Corte de las 22:00 HRS. (272 Secciones)



BERUMEN Y ASOCIADOS, S.C.

GRAFICA 14

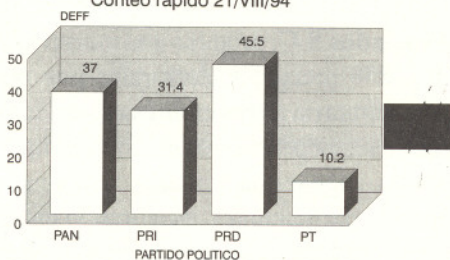
INTERVALOS DE CONFIANZA PPS-PFCRN-PVEM
Corte de las 3:30 Hrs. (481 Secciones)



BERUMEN Y ASOCIADOS, S.C.

GRAFICA 15

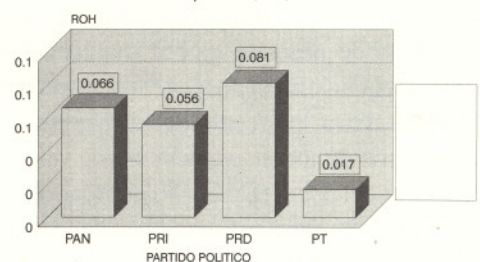
EFFECTOS DE DISEÑO (DEFF)
Conteo rápido 21/VIII/94



BERUMEN Y ASOCIADOS, S.C.

GRAFICA 16

CORRELACION INTRACLASE (ROH)
Conteo rápido 21/VIII/94



BERUMEN Y ASOCIADOS, S.C.

precisión del orden del **1.0%**, y las tres lo mejoraban.

30. En la reunión con las empresas participantes y el Grupo de Asesores Técnicos para Conteo Rápido del **IFE** se acordó que la muestra global sería de **1,500** secciones electorales, dividida en tres réplicas de **500** secciones electorales cada una, mismas que se asignarían en forma aleatoria a cada una de las tres empresas.

Berumen y Asociados S.C. fue la empresa responsable de hacer la selección de muestras para las tres empresas.

31. Para auspiciar la credibilidad del ejercicio y facilitar la explicación ante la sociedad del diseño de muestra a utilizar, se pedía que este fuera autoponderado: el voto válido emitido por cualquier ciudadano “valdría” lo mismo en el cómputo de la estimación que el voto válido emitido por cualquier otro ciudadano. Lo mismo debía ser cierto a nivel sección electoral.

32. Se decidió entonces seleccionar la muestra de **1,500** secciones electorales con un muestreo sistemático aplicado sobre una lista ordenada de secciones electorales.

El orden de la lista fue el siguiente: por entidad federativa, a su interior por número progresivo de Distrito Electoral, a su interior por tipo de sección electoral (urbana, rural, mixta), y a su interior por número consecutivo de la sección electoral (esto aprovecha la estratificación implícita que induce la logística utilizada por el **RFE** para asignar números a las secciones electorales: en “serpentina” o “zig-zag” geográfico.)

33. La muestra de **1,500** secciones electorales se subdividió posteriormente en tres réplicas sistemáticas de **500** secciones electorales cada una.

El procedimiento anterior se aplicó **10** veces generando **10** muestras divididas en tres réplicas cada una. El **IFE** seleccionó al azar una de ellas, asignando aleatoriamente las réplicas correspondientes a cada una de las tres empresas.

V. Errores de muestreo y parámetros asociados

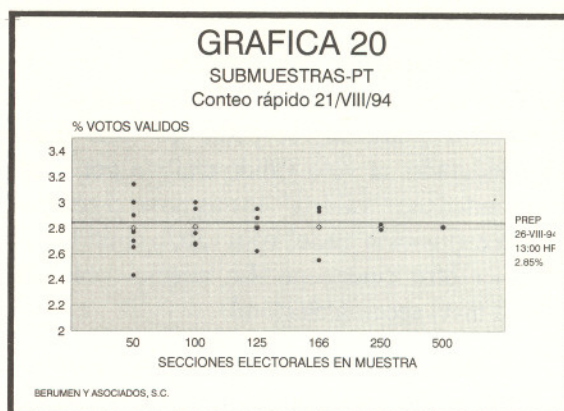
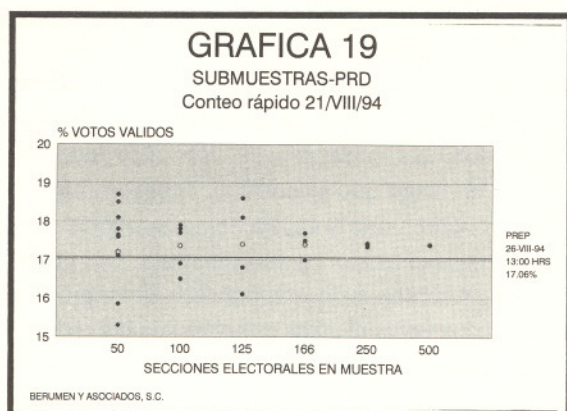
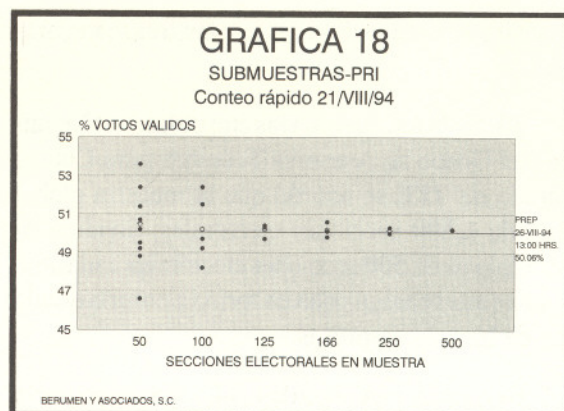
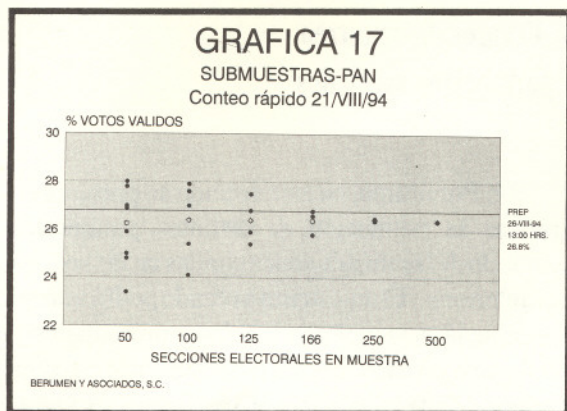
34. Por último, en esta sección se presenta en gráficas la estimación de distintos parámetros resultado de las ilustraciones citadas en las secciones anteriores. El muestrista avezado podrá unir su experiencia personal con los datos presentados y así llegar a diseños futuros más eficientes en estos campos de investigación u otros afines a las variables estimadas.

35. Las gráficas 1-4 presentan las estimaciones a distintos cortes horarios a partir de las **20:00** horas. En ellas se observa la rápida convergencia de las distintas estimaciones. Por otro lado, se aprecia el riesgo que se corre al “comprometer” con el “cliente” una hora fija (temprana) determinada para entregar las estimaciones. Las gráficas 5-8 presentan las correspondientes longitudes medias (dos veces el error estándar) de las estimaciones a los distintos cortes.

La gráfica 9 da la composición urbano-rural mixta de la muestra de secciones electorales, y la composición de las secciones electorales reportadas a los distintos cortes horario; con ello se explica la convergencia de las estimaciones al paso del tiempo cuando la composición de las secciones reportadas se empiezan a asemejar a la de la muestra total.

36. Las gráficas 10-11 presentan las estimaciones, intervalos de confianza, y resultado final del **PREP** para las tres principales fuerzas en la elección de Presidente, para dos cortes de estimación: las **22:00** horas del **21** de agosto y las **3:30** del **22** de agosto. Se observa que la diferencia es mínima y que en ambos cortes, desde el punto de vista estadístico, había la suficiente precisión para declarar el resultado con poca probabilidad de equivocarse en el orden o en la magnitud. Al no intersectarse los intervalos de estimación, el estadístico está en una situación ideal.

37. Para subrayar que esto no siempre es así, en las gráficas 12-14 se presentan los mismos conceptos para los partidos **PPS**, **PFCRN** y **PVEM**, para tres cortes de estimación. Aún en el



último corte de estimación hay un ligero “traslape” entre los intervalos correspondientes a los partidos **PFCRN** y **PVEM**. Desde el punto de vista estadístico, no habría suficiente información para estimar en ese momento cuál de los dos obtuvo más votos.

38. La gráfica 15 presenta los **deff** estimados para cada una de las cuatro principales fuerzas en la elección y la gráfica 16 los **roh** respectivos. Ahí se corrobora que la especulación razonada en la etapa de diseño de utilizar un valor de 46 para **deff** fue acertada. El valor más grande estimado fue de 45.5 para el **PRD**.

39. Por último, para investigar el tamaño de muestra adecuado para obtener estimaciones de

precisión aceptable, que sustenten futuros diseños más eficientes, se particionó la muestra de 500 secciones electorales en 10 submuestras de tamaño 50, cinco de tamaño 100, cuatro de tamaño 125, tres de tamaño 166, y dos de 250.

Con los datos de cada submuestra se obtuvieron las estimaciones respectivas, los resultados se presentan en las gráficas 17-20. Se observa que ya con muestras de tamaño 100 se obtienen estimaciones aceptables, y con tamaño de 166 bastante precisas

Oaxaca, 16-17 de Noviembre de 1994.