



UNIVERSIDAD NACIONAL
AUTONOMA DE MEXICO

Facultad de Ciencias

UNA APLICACION EN ENCUESTAS
ELECTORALES EN EL ESTADO DE
NUEVO LEON

T E S I S

Que para obtener el título de

A C T U A R I O

presenta:

JORGE ALBERTO GODINEZ GARCIA



Director de Tesis: Dra. Guillermina Eslava Gómez

México, D. F.



1999

FACULTAD DE CIENCIAS
SECCION ESCOLAR

TESIS CON
FALLA DE ORIGEN



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



INSTITUTO NACIONAL
DE ESTADÍSTICA Y
CENSO

MAT. MARGARITA ELVIRA CHÁVEZ CANO
Jefa de la División de Estudios Profesionales
P r e s e n t e

Comunicamos a usted que hemos revisado el trabajo de Tesis:

"UNA APLICACION EN ENCUESTAS ELECTORALES EN EL ESTADO DE
NUEVO LEÓN EN 1997"

realizado por GODINEZ GARCIA JORGE ALBERTO

Con número de cuenta 9354713-2 , pasante de la carrera de ACTUARIA

Dicho trabajo cuenta con nuestro voto aprobatorio.

A t e n t a m e n t e

Director de tesis

Propietario

Propietario

Propietario

Suplente

Suplente

DRA. GUILLERMINA ESLAVA GOMEZ

DR. IGNACIO MENDEZ RAMIREZ

MAT. MARGARITA ELVIRA CHAVEZ CANO

ACT. MARIA AURORA VALDES MICHEL

ACT. LAURA MIRIAM QUEROI GONZALEZ

G. Eslava
Jim
México
[Signature]
L. Qu. Q. 9

Consejo Departamental de MATEMATICAS

M. EN A.P. MARIA DEL PILAR ALONSO REYES

DEDICATORIA

Dedico este trabajo con todo mi afecto a:

Mis padres Ramón y Martha,

por ser el pilar principal de mi superación y quienes me dieron las bases de vivir en la perfecta armonía incluyendo el estudio como pieza importante de la superación.

Mis hermanos Mario, Juan, Marco, Ramón y Linda,

para que esto sea de gran motivación hacia ellos.

AGRADECIMIENTOS

Mi novia Georgina, quien ha sacrificado su tiempo para que pueda terminar este trabajo y me ha dado todo su amor sincero.

Guillermina Eslava, que es quien ha dirigido este trabajo y siempre me ha dado sus conocimientos sin condición alguna.

Todos mis profesores, por el tiempo que me han dedicado para poder formarme.

Pablo Macías, Maru Guzmán, Pilar Morfín, Julio Rodríguez, Alejandro Illanes y Eduardo Dueñez, gracias por que con su apoyo pude representar a México en *Olimpiadas Internacionales de Matemáticas*.

A Graciela Blackaller y Rafael Pérez, por ustedes estudié Actuaría.

Wenceslao Hernández, César Morones, Georgina López, Mauricio Morán y Jesús Flores, a quienes les debo la oportunidad de ejercer en el Centro de Estudios de Opinión.

Margarita Chávez, Ignacio Méndez, Aurora Valdés y Laura Querol, por haber revisado este trabajo y dar su voto aprobatorio.

Universidad de Guadalajara y Universidad Nacional Autónoma de México, que son las dos grandes universidades públicas que me han brindado la oportunidad de poder estudiar.

Centro de Estudios de Opinión, por su profesionalismo y visión.

Mi patria, creo en ti, siempre espero servirte y defenderte.

PREFACIO

El objetivo de esta tesis es presentar algunos modelos muestrales que se utilizan en la elaboración de encuestas electorales, la notación seguida es la propuesta por *Sukhatme et al (1984)* principalmente y por *Cochran (1984 versión español)*. La propuesta es mostrar de una manera sencilla y con conocimientos básicos de álgebra y estadística de muestreo aplicado en las encuestas electorales.

Las encuestas electorales en la actualidad son ampliamente difundidas por los medios de comunicación, y son aceptadas en los círculos políticos de México, sobre todo cuando éstas les favorecen.

Como profesionista, una de mis intenciones donde se aplican los conocimientos adquiridos en la Facultad de Ciencias de la Universidad Nacional Autónoma de México y en el Centro de Estudios de Opinión de la Universidad de Guadalajara, es mostrar las metodologías de las encuestas electorales; instrumento joven que se ha convertido en un termómetro político y social dentro de la comunidad mexicana.

Es claro que no sólo existen encuestas electorales o sociales, sino también el muestreo se utiliza para medir "*ratings*" de programación tanto de la radio como de la televisión, para medir los precios al mayoreo y al menudeo, las causas de divorcio, los niveles de popularidad de artistas, la cantidad de fanáticos a favor de un equipo deportivo en un encuentro, el empleo, en los estudios de mercado, en la valuación de cartera morosa, etc. Pero en este trabajo nos concentramos en estudios de opinión políticos y sociales.

Cabe hacer mención que las encuestas en la actualidad tienen algunas herramientas poderosas, la computadora, que ha venido a revolucionar la aplicación del muestreo, si bien es cierto que la teoría ha existido desde hace tiempo, realizarla sin alguna máquina procesadora sería difícil por el tiempo que ésta requiere. En la actualidad, generar tablas de contingencia, por ejemplo, no tiene problema alguno y de igual forma con análisis estadístico.

La descripción de los capítulos de esta tesis se puede encontrar en la introducción que abarca el marco teórico del muestreo; para la elaboración de algunos análisis, se consultaron expertos en la materia, por ejemplo, se contó

con la asesoría del abogado Carlos González para entender el marco jurídico que rodea las encuestas electorales en el estado de Nuevo León, así también colaboraron personas en los estudios de campo para detectar problemas en el levantamiento de las encuestas.

Agradezco las facilidades proporcionadas por el Centro de Estudios de Opinión acerca de los datos de las encuestas de tipo: preelectoral, postvoto, post-electoral y conteo rápido. Asimismo a la Comisión Estatal Electoral de Nuevo León por haberme facilitado los resultados oficiales de la elección para gobernador celebrada el día 6 de julio de 1997 por sección electoral.

Un reconocimiento especial a la Dra. Guillermina Eslava quien estuvo a cargo de la dirección de este trabajo; su trayectoria académica ha sido de gran motivación para mí y en la realización de este trabajo.

También agradezco los comentarios de algunas personas que fueron muy valiosos para poder redactar y mostrar este material.

Jorge Alberto Godínez García

INTRODUCCION

Los cambios democráticos en México se han acelerado últimamente, por lo que las encuestas electorales son una herramienta más de tantas que existen para conocer las tendencias electorales, éstas en México en los últimos seis años se han venido popularizando.

Una base para el diseño de las encuestas electorales es la teoría del muestreo, existen varias obras de muestreo, por ejemplo: *Hansen et al (1951)*, *Kish (1965)*, *Sukhatme et al (1984)*, *Hendricks (1956)*, *Cochran (1977)* y otras. En esta tesis utilizamos particularmente los textos de *Cochran y Sukhatme*.

En el presente trabajo se presenta la aplicación del muestreo en encuestas electorales realizadas en el estado de Nuevo León, en la elección para gobernador realizada el día 6 de julio de 1997. Desde una encuesta preelectoral hasta una encuesta post-electoral, de alguna manera estamos en condiciones de presentar un ciclo completo que comienza con una encuesta preelectoral, le sigue una encuesta de salida o postvoto, el conteo rápido, la publicación de los resultados oficiales por el órgano legal para hacerlo y por último la encuesta post-electoral. ¿Qué es cada una de estas encuestas?, ¿Para qué sirven?, ¿Cuál es el objetivo de cada una?, son cuestionamientos que se explican en la tesis para poder comprender cada una de las encuestas. Se pueden ver gráficamente los resultados de cada una de las encuestas electorales, el conteo rápido y los resultados oficiales, además de una breve explicación de cada una de ellas.

En el primer Capítulo se pueden ver algunos datos sociodemográficos y electorales del estado de Nuevo León, con la idea de complementar el conocimiento hacia el universo de estudio. Para la elaboración de un diseño muestral es muy importante tener información sobre el universo a estudiar. Se hace referencia al marco legal en el apartado 1.7, este es un punto importante ya que en la actualidad existen leyes que regulan la publicación de encuestas electorales, cada estado tiene su propio marco legal y es necesario que la metodología que se da a conocer a la luz pública, cumpla con los requisitos de éste.

El soporte teórico de las encuestas, como se ha mencionado, es la teoría del muestreo, por consiguiente en el Capítulo dos se mostrarán los diseños del muestreo como son: aleatorio simple (explicado en el apartado 2.1), éste es el muestreo probabilístico básico que existe en la teoría; también se presenta el

estratificado y sus características (se presenta en el apartado 2.2); el muestreo por conglomerados; el muestreo bietápico o submuestreo; el submuestreo estratificado que es la interacción entre el muestreo estratificado y el muestreo de dos etapas; el muestreo trietápico; y el muestreo por cuotas que son utilizadas en algunas ocasiones.

En el Capítulo tres se analizan las metodologías empleadas en las encuestas aplicadas en el estado de Nuevo León, las metodologías son acordes a la teoría mostrada en el Capítulo dos y también se presenta la definición de cada proceso electoral.

En el Capítulo cuatro se muestran las principales gráficas de cada una de las encuestas y una breve explicación, esto es para entender mejor una encuesta electoral y la evolución de la misma. En el apartado 4.7 se incluyen los cuadros comparativos de cada una de las encuestas para conocer la precisión de las mismas respecto al resultado oficial, pero también es importante conocer que una encuesta es sólo un reflejo de lo que ocurre en el universo de estudio al momento del levantamiento de ésta, ya que existen fenómenos sociales que hacen cambiar de opinión a la sociedad de un momento a otro, un ejemplo podría ser los debates políticos, en la actualidad éstos tienen importancia sobre el electorado, mucho del futuro político de cada uno de los partidos políticos está en un solo acontecimiento, esta es una variable social que puede modificar los resultados de las encuestas rápidamente.

En el Capítulo cinco se presentan breves conclusiones y comentarios de lo que se ha expuesto en esta tesis, la importancia de las encuestas en la actualidad y la visión de las mismas hacia el futuro.

INDICE

Prefacio	i
Introducción	iii
Índice	v
CAPÍTULO 1 Datos Sociodemográficos y electorales en el estado de Nuevo León	1
1.1 Población total	1
1.2 Distribución geográfica	2
1.3 Composición por edad y género	2
1.4 Migración	2
1.5 Alfabetismo	2
1.6 Nivel de instrucción	3
1.7 Marco legal	3
1.8 Población electoral	5
1.9 Distribución geográfico electoral	5
CAPÍTULO 2 Algunos diseños de muestreo	10
Introducción	10
2.1 Muestreo aleatorio simple	10
2.2 Muestreo aleatorio estratificado	22
2.3 Muestreo por conglomerados	27
2.4 Submuestreo o muestreo bietápico	31
2.5 Submuestreo estratificado	37
2.6 Muestreo trietápico	40
2.7 Efecto de diseño	47
2.8 Muestreo por cuotas	49
CAPÍTULO 3 Diseños muestrales de las encuestas preelectoral, de salida o postvoto, post-electoral y el conteo rápido en el estado de Nuevo León	54
3.1 Introducción	54
3.2 Encuesta preelectoral	54

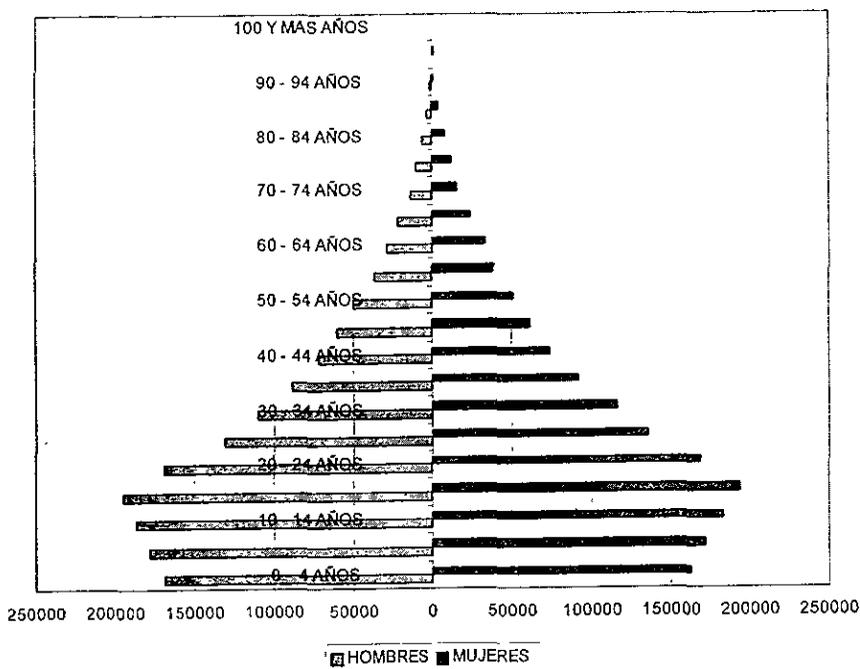
3.3 Encuesta de salida o postvoto	63
3.4 Conteo rápido	69
3.5 Encuesta Post-electoral	73
CAPÍTULO 4 Presentación de resultados	80
4.1 Introducción	80
4.2 Encuesta preelectoral	80
4.3 Encuesta de salida o postvoto	101
4.4 Conteo rápido	110
4.5 Resultados oficiales	120
4.6 Encuesta post-electoral	121
CAPÍTULO 5 Conclusiones y comentarios	138
5.1 De los procesos electorales	138
5.2 Resultados comparativos	141
5.3 Conclusiones y comentarios globales	151
BIBLIOGRAFÍA	155

1.- DATOS SOCIODEMOGRAFICOS Y ELECTORALES DEL ESTADO DE NUEVO LEON

1.1 POBLACION TOTAL

La población en el año de 1990 según el censo del 90 hecho por el Inegi y publicado en el código90 era de 3'098,736 habitantes, en el año de 1995 según el Inegi y publicado en el anuario estadístico del estado de Nuevo León edición 1996 del conteo95 la población total era de 3'549,273. La tasa de crecimiento anual del estado es del 2.76% de acuerdo al censo del 90 y al conteo realizado el 5 de Noviembre de 1995, menor al 2.81% de la tasa nacional.

Gráfico 1-1 1 Pirámide poblacional del estado de Nuevo León según el código90 publicado por el Inegi



1.2 DISTRIBUCION GEOGRAFICA.

De los 51 municipios que tiene el estado, el municipio más poblado es el de Monterrey con 1'069,238 habitantes que representa el 34.51%, asimismo Guadalupe con 535,560 habitantes que representan el 17.28% y San Nicolás de los Garza concentra 436,603 habitantes que representan el 14.09% del total del estado de Nuevo León, de esta manera estos municipios son los más poblados del estado.

Las cifras de población presentadas consideran a los residentes en viviendas particulares y colectivas, al personal diplomático que radica en el extranjero así como a su familia y dependientes económicos y la población sin vivienda que vive en la vía pública.

Cabe hacer mención que del total de la población, se divide en localidades urbanas y rurales, según el Inegi una localidad rural es aquella que tiene menos de 2,500 habitantes y urbana es aquella que tiene 2,500 ó más o es cabecera municipal, con esta clasificación el 7.29% del estado es población rural, mientras que el 92.71% habita en zona urbana.

1.3 COMPOSICION POR GENERO

La población tiene una estructura por género en la que el masculino alcanza la minoría, siendo una cantidad de 1'542,551 hombres que representan el 49.78% y 1'556,185 mujeres que representan el otro 50.22%.

1.4 MIGRACION

Con respecto a la migración, se observa que el 23.11% de los habitantes no es nacido en la entidad y el 76.89% es nacido en la entidad.

1.5 ALFABETISMO

El 95.04% de la población mayor de 6 años es alfabeto, por consiguiente el otro 4.96% de la población mayor de 6 años es analfabeto en el estado de Nuevo León; la proporción de analfabetos en Nuevo León es menor al 7.26% en el ámbito nacional.

1.6 NIVEL DE INSTRUCCION

En el ámbito de nivel de instrucción, la población de 6 años y más del estado de Nuevo León se puede dividir de la siguiente forma: el 7.80% no tiene instrucción primaria, el 3.98% tiene el primer grado, el 5.21% el segundo grado, el 7.00% tiene el tercer grado, el 5.88% tiene el cuarto grado, el 4.69% tiene el quinto grado, el 17.21% el sexto grado, el 46.88% tiene post-primaria y el 1.35% no está especificado.

1.7 MARCO LEGAL

La Ley Electoral Del Estado De Nuevo León, es el órgano que regula el ejercicio de los derechos políticos de los ciudadanos del estado, la constitución y funcionamiento de las organizaciones políticas; la preparación, desarrollo y vigilancia del proceso electoral para la renovación de los Poderes Legislativo y Ejecutivo y de los Ayuntamientos del estado; la declaración de los resultados electorales y la resolución de lo contencioso electoral como dice el primero de los 295 artículos y 10 transitorios en un capítulo único que la conforman, además, la Ley contiene tres partes, la primera parte tiene cinco capítulos, la segunda parte tiene tres capítulos y la tercera parte tiene tres capítulos; la primera parte habla del objeto de la ley de las organizaciones electorales, la segunda parte del proceso electoral y la tercera parte de lo contencioso electoral. La mencionada Ley fue promulgada el 13 de diciembre de 1996 y presentamos los puntos más importantes referente al tema de las encuestas electorales.

El artículo 139 menciona que aquellos que soliciten u ordenen la publicación de cualquier encuesta o sondeo de opinión sobre asuntos electorales que se realice desde el inicio de las campañas hasta el cierre oficial de las casillas (las 18:00hrs) el día de la elección, deberá entregar copia del estudio completo al Presidente de la Comisión Estatal Electoral, si la encuesta o sondeo se difundirá por cualquier medio. En todo caso, la difusión de los resultados de cualquier encuesta o sondeo de opinión, estará sujeta a lo dispuesto en el párrafo siguiente.

Durante los ocho días previos a la elección y hasta la hora del cierre de las casillas queda prohibido publicar o difundir, por cualquier medio, los

resultados de encuestas o sondeos de opinión que tengan por objeto dar a conocer las preferencias electorales de los ciudadanos.

Asimismo existe un acuerdo de la Comisión Electoral para vigilar el cumplimiento de las disposiciones constitucionales contenidas en el Código.

Primero: Quienes realicen estudios relacionados con sondeos de opinión, encuestas y otros de carácter estadístico relacionados con el proceso electoral de renovación del poder Ejecutivo del Estado y difundan los resultados deberán sujetarse a los siguientes criterios y normas.

1) Proporcionar a la opinión pública la llamada vitrina metodológica en la que basan sus estudios detallando los siguientes aspectos:

- Tipo de muestreo
- Marco muestral (un requisito mínimo para pertenecer al marco muestral debe ser, el tener credencial de elector actualizada)
- Diseño de muestra
- Tamaño de muestra
- Selección de la muestra
- Márgenes de error y validación estadísticas
- Tipo de trabajo de campo (lugares, fechas y forma del levantamiento de la encuesta)
- Formas y procedimientos para el proceso de información
- También deberá indicarse si los resultados son completos o sólo parte de la información obtenida

2) Otros aspectos que garanticen la credibilidad y confianza en el estudio y sus resultados, tales como datos de identificación de la empresa o institución encuestadora así como del patrocinador de la misma.

Segundo: Los resultados y todos los aspectos señalados en el acuerdo primero que tengan que ver con sondeos, encuestas y estudios estadístico-electorales relacionados con aspectos previos a la elección podrán difundirse durante el tiempo que dure el proceso y hasta las veinticuatro horas del día 29 de junio de 1997. De igual forma queda prohibida su publicación durante la propia jornada electoral.

Tercero: Todos los estudios de encuesta de salida que se realicen además de cumplir con los criterios señalados en el primer punto de este acuerdo podrán ser dados a conocer dos horas después de la clausura de las casillas establecida por el Código.

Cuarto: La Comisión Estatal Electoral está en la disposición de aprobar la metodología empleada en la realización de la encuesta electoral siempre y cuando la persona física, empresa, grupo, organismo o asociación que desee validar dicha metodología, cumpla con todos los requisitos establecidos en el presente acuerdo. Para ello deberán proporcionar una copia completa del estudio ante este organismo electoral para su revisión.

Es importante hacer mención que de no cumplirse los requisitos no existe sanción alguna.

1.8 POBLACION ELECTORAL

La población de 18 años o más constituye el 57.90% de la población total del estado, de los cuales se registró un total en lista nominal de 2'161,453. El estado tiene 2,123 secciones electorales, la información que se menciona fue proporcionada por el Instituto Federal Electoral al corte del 6 de julio de 1997.

1.9 DISTRIBUCION GEOGRAFICO ELECTORAL

En la elección del 6 de julio para elegir gobernador del estado de Nuevo León se tienen 11 Distritos Electorales Federales y 26 Distritos Electorales Estatales. La distribución de la lista nominal por municipio, por distrito electoral estatal y distrito electoral federal se presentan en los cuadros 1-1, 1-2 y 1-3:

Cuadro 1-1: Lista nominal de la elección para gobernador del estado de Nuevo León celebrada el día 6 de julio de 1997, información proporcionada por municipio.

MUNICIPIO	LISTA NOMINAL	%	MUNICIPIO	LISTA NOMINAL	%
ABASOLO	1,201	0.06%	HERRERAS, LOS	2,430	0.11%
AGUALEGUAS	3,020	0.14%	HIDALGO	7,701	0.36%
ALDAMAS, LOS	1,972	0.09%	HIGUERAS	1,323	0.06%
ALLENDE	15,686	0.73%	HUALAHUISES	3,725	0.17%
ANAHUAC	10,727	0.50%	ITURBIDE	2,385	0.11%
APODACA	108,063	5.00%	LAMPAZOS DE NARANJO	2,529	0.12%
ARAMBERRI	8,109	0.38%	LINARES	40,259	1.86%
BENITO JUAREZ	20,285	0.94%	MARIN	1,995	0.09%
BUSTAMANTE	2,158	0.10%	MELCHOR OCAMPO	929	0.04%
CADEREYTA JIMENEZ	37,046	1.71%	MIER Y NORIEGA	3,942	0.18%
CARMEN	3,470	0.16%	MINA	2,567	0.12%
CERRALVO	4,877	0.23%	MONTEMORELOS	32,268	1.49%
CIENEGA DE FLORES	4,914	0.23%	MONTERREY	720,758	33.35%
CHINA	7,168	0.33%	PARAS	1,334	0.06%
DOCTOR ARROYO	18,593	0.86%	PESQUERIA	5,825	0.27%
DOCTOR COSS	1,993	0.09%	RAMONES, LOS	4,014	0.19%
DOCTOR GONZALEZ	1,670	0.08%	RAYONES	1,832	0.08%
GALEANA	22,469	1.04%	SABINAS HIDALGO	19,562	0.91%
GARCIA	11,265	0.52%	SALINAS VICTORIA	8,355	0.39%
GENERAL BRAVO	3,612	0.17%	SAN NICOLAS DE LOS GARZA	304,160	14.07%
GENERAL ESCOBEDO	87,348	4.04%	SAN PEDRO GARZA GARCIA	78,460	3.63%
GENERAL TERAN	9,977	0.46%	SANTA CATARINA	116,539	5.39%
GENERAL TREVIÑO	1,223	0.06%	SANTIAGO	21,063	0.97%
GENERAL ZARAGOZA	2,718	0.13%	VALLECILLO	2,571	0.12%
GENERAL ZUAZUA	3,334	0.15%	VILLALDANA	2,709	0.13%
GUADALUPE	379,350	17.55%	TOTAL	2,161,453	100.00%

Cuadro 1-2: Lista nominal de la elección para gobernador del estado de Nuevo León celebrada el día 6 de julio de 1997, información proporcionada por Distrito Local Electoral

DISTRITO	LISTA NOMINAL	%	DISTRITO	LISTA NOMINAL	%
1	105,800	4.89%	14	104,839	4.85%
2	95,039	4.40%	15	89,750	4.15%
3	97,861	4.53%	16	108,063	5.00%
4	86,235	3.99%	17	101,259	4.68%
5	83,086	3.84%	18	97,139	4.49%
6	81,005	3.75%	19	97,860	4.53%
7	83,240	3.85%	20	44,130	2.04%
8	88,492	4.09%	21	40,256	1.86%
9	97,690	4.52%	22	62,317	2.88%
10	93,409	4.32%	23	58,109	2.69%
11	99,150	4.59%	24	57,931	2.68%
12	97,291	4.50%	25	45,816	2.12%
13	87,470	4.05%	26	58,216	2.69%
			TOTAL	2,161,453	

DISTRITOS 1, 2, 3, 4, 5, 6, 7, 8 MONTERREY	DISTRITOS 9, 10, 11 SAN NICOLAS DE LOS GARZA	DISTRITOS 12, 13, 14, 15 GUADALUPE	DISTRITO 16 APODACA
DISTRITO 17 GENERAL ESCOBEDO SAN NICOLAS DE LOS GARZA	DISTRITO 18 SAN PEDRO GARZA GARCIA SANTA CATARINA	DISTRITO 19 SANTA CATARINA	DISTRITO 20 ABASOLO CARMEN CIENEGA DE FLORES GARCIA GENERAL ZUAZUA HIDALGO HIGUERAS MINA SALINAS VICTORIA
DISTRITO 21 ANAHUAC BUSTAMANTE LAMPAZOS DE NARANJO SABINAS HIDALGO VALLECILLO VILLALDANA	DISTRITO 22 AGUALEGUAS ALDAMAS, LOS BENITO JUAREZ CERRALVO CHINA DOCTOR COSS DOCTOR GONZALEZ GENERAL BRAVO GENERAL TREVIÑO HERRERAS, LOS MARIN MELCHOR OCAMPO PARAS PESQUERIA RAMONES, LOS	DISTRITO 23 CADEREYTA JIMENEZ SANTIAGO DISTRITO 24 ALLENDE GENERAL TERAN MONTEMORELOS DISTRITO 25 HUALAHUISES LINARES RAYONES	DISTRITO 26 ARAMBERRI DOCTOR ARROYO GALEANA GENERAL ZARAGOZA ITURBIDE MIER Y NORIEGA

Cuadro 1-3: Lista nominal de la elección para gobernador del estado de Nuevo León celebrada el 6 de julio de 1997, información proporcionada por Distrito Federal Electoral

DISTRITO ELECTORAL FEDERAL	LISTA NOMINAL	%	DISTRITO ELECTORAL FEDERAL	LISTA NOMINAL	%
1	224,245	10.37%	7	185,895	8.60%
2	242,311	11.21%	8	183,221	8.48%
3	192,801	8.91%	9	203,281	9.40%
4	198,907	9.20%	10	171,732	7.95%
5	191,425	8.86%	11	196,129	9.07%
6	171,706	7.94%	TOTAL	2,161,453	100.00%

DISTRITO 1
ANAHUAC
BUSTAMANTE
GARCIA
LAMPAZOS DE NARANJO
MINA
SAN PEDRO GARZA GARCIA
SANTA CATARINA

DISTRITO 4
SAN NICOLAS DE LOS GARZA

DISTRITO 6
MONTERREY

DISTRITO 8
GUADALUPE

DISTRITO 10
MONTERREY

DISTRITO 11
GUADALUPE

DISTRITO 2
ABASOLO
AGUALEGUAS
ALDAMAS, LOS
APODACA
CADEREYTA, JIMENEZ
CARMEN
CERRALVO
CIENEGA DE FLORES
CHINA
DOCTOR COSS
DOCTOR GONZALEZ
GENERAL BRAVO
GENERAL TREVIÑO
GENERAL ZUAZUA
HERRERAS, LOS
HIDALGO
HIGUERAS
MARIN
MELCHOR OCAMPO
PARAS
PESQUERIA
RAMONES, LOS
SABINAS HIDALGO
SALINAS VICTORIA
VALLECILLO
VILLALDANA

DISTRITO 3
GENERAL ESCOBEDO
SAN NICOLAS DE LOS GARZA

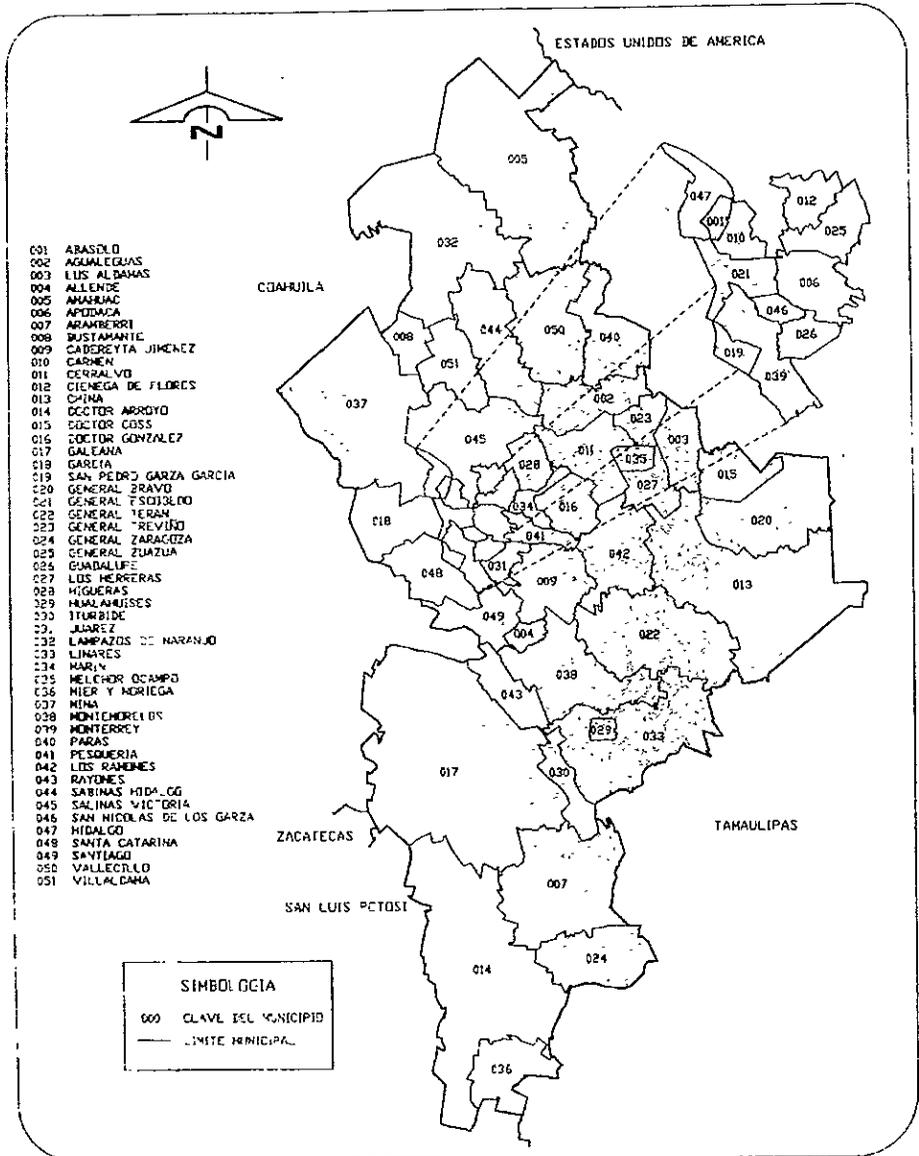
DISTRITO 5
MONTERREY

DISTRITO 7
MONTERREY

DISTRITO 9
ALLENDE
ARAMBERRI
BENITO JUAREZ
DOCTOR ARROYO
GALEANA
GENERAL TERAN
GENERAL ZARAGOZA
HUALAHUISES
ITURBIDE
LINARES
MIER Y NORIEGA
MONTEMORELOS
RAYONES
SANTIAGO

Mapa 1-A, del estado de Nuevo León con división municipal.

División Geoestadística Municipal



FUENTE: INEGI, Marco Geoestadístico, 1995. Inédito

2.- ALGUNOS DISEÑOS DE MUESTREO

INTRODUCCION

En este capítulo se presentan algunos diseños de muestreo como es el Muestreo Aleatorio Simple (MAS), se analizará un ejemplo, después se verá el muestreo aleatorio estratificado, su definición y algunas de sus aplicaciones, el tamaño de muestra se puede ver en el apartado 2.1.4, en el apartado 2.1.5 se verán estimadores para proporciones, posterior a esto se verá el muestreo por conglomerados, las ventajas del mismo, los costos y los estimadores, después se verá el submuestreo o muestreo bietápico y se presenta el estimador trietápico, además se presenta el submuestreo estratificado, el efecto de diseño y el muestreo por cuotas.

2.1. MUESTREO ALEATORIO SIMPLE

2.1.1 INTRODUCCION

El Muestreo Aleatorio Simple (MAS) es un método de selección de n unidades en un conjunto de N , de esta manera existen ${}_N C_n$ (donde ${}_N C_n$ son las combinaciones de N elementos en n) muestras distintas que tienen la misma probabilidad de ser elegidas, además cada uno de los elementos tiene la misma probabilidad de ser seleccionado en cualquier extracción, $1/N$.

$${}_N C_n = \frac{N!}{n!(N-n)!}$$

Ejemplo: una persona de dos que quieren ser presidente de una escuela, dice a su amigo ser el claro triunfador si las elecciones fueran en este momento, su amigo le dice que sí, de hecho le muestra a sus amigos y le dicen que van a votar por él, pero uno se hace la pregunta obligada, ¿acaso al igual que los demás candidatos todos dicen ser los triunfadores? y ¿no subestimaré a la fuerza opositora?; entonces alguien de por allí le propone que se entrevisten a quince personas pero no de los amigos, sino que sean personas que encuentren al azar, el candidato inmediatamente le dice que es una persona desconfiada, que le extraña que dude de su buena fe, pero para demostrarle que él (candidato) tiene la razón acepta el reto, se sabe que la escuela tiene 2,000 estudiantes.

Desde luego que esto no es de nuestro pleno interés, mas bien serían las siguientes preguntas: ¿Que tan representativo son 15 encuestas entre los 2,000 estudiantes?, ¿Que decisión tomaría la persona si sólo 5 votarían a favor de él de los 15 entrevistados?

Esta persona tiene una noción del azar, pero en realidad lo que está haciendo sin saberlo es una aplicación muy cercana al MAS.

2.1.2. APLICACION DEL MUESTREO ALEATORIO SIMPLE

Si en realidad la persona deseara aplicar un método probabilístico de MAS seguiría los siguientes pasos:

1. - Asignar un número del 1 al 2,000 a cada uno de los estudiantes.
2. - Acomodarlos de una manera práctica, de tal manera que sea sencillo identificarlos en caso dado de ser seleccionados.
3. - Se producen 15 números aleatorios, que pueden ser usando la computadora, que son números del intervalo $(0,1)$, pues los algoritmos implementados en la computadora generalmente produce números aleatorios entre 0 y 1; se multiplica por 2,000 el número aleatorio y nos genera un número aleatorio que nos servirá de base para la selección de los 15 estudiantes, de esta manera se pueden elegir a los estudiantes.

La computadora ha venido a facilitar mucho el trabajo del muestreo, la generación de números aleatorios es uno de ellos, las tablas de números aleatorios son igualmente válidas en este proceso, pero el tiempo que nos lleva la computadora comparativamente con la tabla es muy diferente, sobre todo en la selección de unidades de muestreo con diferente probabilidad de ser seleccionadas.

2.1.3. NOTACION DEL MUESTREO ALEATORIO SIMPLE

La notación del MAS que aquí se presenta es la que aparece en *Sukhatme et al, (1984, págs. 25 - 29)*.

Sea Y el atributo que está sujeto a medición, denotamos por Y_i el valor del atributo asociado con la unidad U_i , donde $i = 1, 2, 3, \dots, N$.

La media poblacional

$$\bar{Y} = \frac{1}{N} \sum_{i=1}^N Y_i, \quad (2.1.1)$$

cuadrados medios de la población

$$S^2 = \frac{1}{N-1} \sum_{i=1}^N (Y_i - \bar{Y})^2 = \frac{1}{N-1} \left(\sum_{i=1}^N Y_i^2 - NY^2 \right), \quad (2.1.2)$$

población total

$$T(Y) = \sum_{i=1}^N Y_i, \quad (2.1.3)$$

varianza poblacional

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (Y_i - \bar{Y})^2 = \frac{N-1}{N} S^2, \quad (2.1.4)$$

$f = n/N$ se conoce como fracción de muestreo

donde y es (2.1.6).

$$V(\bar{y}) = \left(\frac{N-n}{N} \right) \frac{S^2}{n} = (1-f) \frac{S^2}{n}, \quad (2.1.5)$$

2.1.3. ESTIMADORES DEL MUESTREO ALEATORIO SIMPLE

Estimador insesgado de la media muestral

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i, \quad (2.1.6)$$

cuadrados medios de la muestra

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2, \quad (2.1.7)$$

estimador de la varianza de la media muestral

$$\hat{v}(\bar{y}) = (1-f) \frac{s^2}{n}, \quad (2.1.8)$$

error estándar de \bar{y}

$$\sigma_{\bar{y}} = \sqrt{\left(\frac{N-n}{N}\right) \left(\frac{S^2}{n}\right)} = (1-f) \frac{S}{\sqrt{n}}. \quad (2.1.9)$$

2.1.4 TAMAÑO DE LA MUESTRA

Un punto importante después o antes del diseño muestral es el tamaño de muestra, según *Cochran (1977, págs. 73-74)* los pasos principales involucrados en la elección del tamaño de la muestra son los siguientes:

1.- Debe existir algún enunciado respecto a lo que se espera de la muestra. Este puede darse en términos de límites de error deseados, o bien en términos de alguna decisión o acción que debe tomarse una vez que se conocen los resultados de la muestra. La responsabilidad de este enunciado es primordialmente de las personas que van a usar los resultados de la encuesta, aunque con frecuencia, dichas personas necesitan una guía para expresar sus deseos en términos numéricos.

2.- Se debe encontrar una ecuación que relacione n con la precisión deseada de la muestra. La ecuación variará según el contenido del enunciado de precisión y el tipo de muestreo propuesto. Una de las ventajas del muestreo probabilístico es que existe una fórmula explícita para el cálculo del tamaño de muestra.

3.- Esta fórmula tendrá como parámetros ciertas propiedades desconocidas de la población, que deben estimarse para obtener resultados específicos.

4.- Con frecuencia sucede que los datos estipulan para ciertas subdivisiones mayores de la población y que los límites de error deseados se establecen para cada subdivisión o dominios de estudio. De ser así, se hace un cálculo separado para el valor n en cada subdivisión y el n total se encuentra por adición.

5.- Generalmente se mide más de un atributo o característica en una encuesta por muestreo: en ocasiones, el número de atributos es grande. Si se define un grado de precisión para cada atributo, debe encontrarse un método para reconciliar estos valores, una posibilidad sería tomar el tamaño de muestra máximo de todos los atributos, otra forma sería tomar los principales atributos y tomar el tamaño de muestra máximo.

6.- Finalmente, debe apreciarse el valor elegido de n , para que sea consistente con los recursos del muestreo disponibles. Esto exige una estimación del costo, trabajo, tiempo y materiales que se necesitan para obtener la muestra del tamaño propuesto. En ocasiones es claro que n debe reducirse drásticamente, y entonces es necesario tomar una decisión, que es la de proceder con una muestra mucho más pequeña, lo que reduce la precisión, o bien, abandonar los esfuerzos hasta contar con mayores recursos.

El tamaño de muestra debe considerar los costos y la precisión, la muestra debe por definición obtener la precisión requerida al menor costo. Si el tamaño de muestra es muy pequeño e insuficiente se pierde precisión, si por el otro lado es muy grande, los costos del levantamiento son muy costosos y la precisión ganada en el estudio es mínima, eso de la precisión es relativo, la realidad es que cuando se hacen tamaños de muestra muy grandes para fines de obtener información más detallada (cruces de variables) se pierde un poco de vista el objetivo principal y eso lleva en la práctica a errores en el levantamiento por la gran cantidad de personas que se deben manejar.

Una manera de estimar el tamaño de muestra es conociendo la varianza o un estimador de la varianza del universo de estudio mediante un estudio inmediato anterior o una encuesta piloto, desafortunadamente casi siempre no se tiene información al respecto, de esa manera no es sencillo obtener un tamaño de muestra de manera “natural”, por consiguiente, debe uno de apoyarse en la experiencia y tomar una decisión.

Para dejar más claro esto, vamos a poner un ejemplo, regresando al ejemplo del candidato, él afirma que obtendrá el 60% de los votos a su favor y por lo tanto será ganador, nosotros conocemos que el tamaño del universo de estudio es 2,000 de individuos, la muestra requiere que sea representativa del universo de estudio, sabemos además, que las unidades de selección son los individuos de la escuela y que contestan con la verdad, necesitamos que los resultados de la encuesta sean confiables, después de esto le mencionamos que el parámetro le arrojará el resultado del universo de estudio al momento del levantamiento de la encuesta.

Ahora ayudemos al candidato a que nos explique sus requerimientos en números, para esto le mencionamos que existen dos puntos que se deben tomar en cuenta: la precisión que se puede establecer como la cantidad de error tolerable en las estimaciones muestrales; en la realidad es difícil determinar la cantidad de error tolerable, para esto le pedimos al candidato que nivel de error está él dispuesto a tolerar, desde luego que antes debemos de saber cual es el parámetro que deseamos probar: digamos que el parámetro que deseamos probar es el porcentaje de alumnos que votarían por el candidato que él sospecha está cerca del 60%. El candidato nos contesta que está dispuesto a tolerar un error de un $\pm 5\%$, para el candidato un límite de error de $\pm 5\%$ lo considera lo suficientemente pequeño para permitir clasificar a los alumnos que votarían por él y los que no votarían por él. Prácticamente de esta manera es como se suelen resolver los límites de error, cuando por primera vez se pregunta a las personas el grado de precisión deseado, confiesan que la mayoría de las veces dicen no tener idea de lo que se les está hablando, pero después de tener una charla en la que el experto trata de indicarle lo que significa *Cochran (1984, pág 106)*, mencionan que se dan cuenta gracias a su experiencia, que después de discutir el tema, son capaces de indicar el grado de precisión deseado, al menos aproximadamente el tamaño del límite de error que les parece razonable. Pero casi siempre existe un cierto nivel de adivinanza en la especificación de la precisión para

un futuro más o menos inmediato. La otra cuestión sería que nos diese la probabilidad sugerida para que nuestro parámetro quede dentro de la tolerancia, a lo que el candidato nos contesta que desea una probabilidad de que el parámetro caiga dentro de la tolerancia de 95%.

$$P(p - p \leq d) = 1 - \alpha$$

Donde P es la proporción poblacional no conocida y p es la proporción obtenida de la muestra. Es posible determinar el tamaño de muestra que cumpla con los requerimientos del candidato, el tamaño lo calcularemos en dos procesos, uno calculando n_0 y otro calculando n sobre la base de n_0 en términos de p y q , donde p y q son obtenidos de una encuesta piloto o casos previos.

$$n_0 = \frac{t^2 pq}{d^2} \quad (2.1.10)$$

Donde d es el error absoluto, t es la abscisa de la curva normal que corta un área de $\alpha = 1 - 0.95 = 0.05$, en las colas de una distribución normal, p es el porcentaje de votantes a favor del candidato y $q = 1 - p$, \therefore se tiene:

$$t = 1.96 \quad d = \pm 0.05 \quad p = 0.60 \quad \Rightarrow$$

$$n_0 = \frac{(1.96)^2 (0.6) (0.4)}{(0.05)^2} = \frac{0.9219}{0.0025} = 368.79 \cong 369.$$

Queda claro, que en este primer proceso no utilizamos el tamaño del universo, este tamaño de muestra n_0 es ligeramente más grande que n , n_0 supone que el tamaño del universo es infinito, la fórmula del segundo proceso es el siguiente:

$$n = \frac{n_0}{1 + \frac{n_0 - 1}{N}} \quad (2.1.11)$$

sustituyendo $n_0 = 369$ y $N = 2,000$ nos queda el siguiente resultado:

$$n = \frac{369}{1 + \frac{(369-1)}{2,000}} = 311.65 \cong 312.$$

La diferencia es de 57, lo que generalmente genera un ahorro de costos y algo muy importante en las encuestas de la actualidad, el tiempo.

Existe a su vez un método para calcular el tamaño de muestra para proporciones pequeñas, hablar de un estimador de P alrededor del 5% con un error absoluto de $\pm 5\%$ es aventurado, generalmente para $P \leq 20\%$ se utiliza el error relativo, es decir, se utiliza el coeficiente de variación en lugar de la varianza.

Así $r =$ error relativo

$$P(p - P \geq rP) = \alpha,$$

donde rP sustituye a d .

Por consiguiente queda la siguiente fórmula:

$$n_0 = \frac{t^2 pq}{r^2 p^2} = \frac{t^2 q}{r^2 p}, \quad (2.1.12)$$

y sustituyendo n_0 de (2.1.12) en (2.2.11) se obtiene el tamaño de muestra para poblaciones con proporciones pequeñas.

2.1.5. ESTIMADORES PARA PROPORCIONES DEL MUESTREO ALEATORIO SIMPLE

En ciertas ocasiones deseamos estimar el porcentaje o proporción de unidades de la población que poseen cierta característica o atributo, o que caen dentro de alguna clase definida. La mayoría de los resultados de las encuestas de opinión pública son de este tipo, basta tomar de ejemplo la intención de voto de los ciudadanos de un universo de estudio, el porcentaje de ciudadanos que votan por el partido X, la proporción de personas que están a favor de una política exterior determinada. La información puede ser en forma directa dentro del cuestionario y generalmente cerrada, la función que se determina es dicotómica, es decir, se puede contestar con un sí con un valor de 1, en otro caso diferente a "sí" vale 0.

Supongamos que todas y cada una de las unidades en la población caen dentro de una y sólo una de las dos posibles clases C y C'. La notación que aquí se usa es la de *Cochran (1984, págs 80-81)* ver cuadro 2-1.

Cuadro 2-1: Cuadro diferencial entre proporciones y total de una muestra y su población.

Número de unidades en C en		Proporción de unidades en C en	
Población	Muestra	Población	Muestra
A	a	$P = A/N$	$p = a/n$

La estimación muestral de P es p , y la estimación muestral de A es Np o Na/n .

Para calcular la varianza se puede aplicar la fórmula (2.1.8) a esta situación. Para cualquier unidad en la muestra o población, se define y_i como 1 si la unidad se encuentra en C, y como 0 si está en C'.

$$Y = \sum_{i=1}^N y_i = A, \quad (2.1.13)$$

$$\bar{Y} = \sum_{i=1}^N y_i / N = A / N = P. \quad (2.1.14)$$

En la misma forma para el estimador de la media

$$\bar{y} = \sum_{i=1}^n y_i / n = a / n = p. \quad (2.1.15)$$

De esta manera estimar A y P es similar a la estimación del total y la media de una población en la cual, todos los valores y_i son 1 ó 0, primero expresemos S^2 y s^2 en términos de P y p . Debemos notar que:

$$\sum_{i=1}^N y_i^2 = A = NP, \quad \sum_{i=1}^n y_i^2 = a = np,$$

$$S^2 = \frac{\sum_{i=1}^N (y_i - \bar{Y})^2}{N-1} = \frac{\sum_{i=1}^N y_i^2 - N\bar{Y}^2}{N-1} = \frac{1}{N-1} (NP - NP^2) = \frac{N}{N-1} PQ. \quad (2.1.16)$$

Donde $Q = 1 - P$. De manera similar,

$$s^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1} = \frac{n}{n-1} pq. \quad (2.1.17)$$

Para un muestreo aleatorio simple de las unidades clasificadas C y C' tenemos que $p = a/n$ es un estimador insesgado de $P = A/N$ y la varianza de p usando (2.1.15) es:

$$V(p) = E(p - P)^2 = \left(\frac{N-n}{N} \right) \left(\frac{S^2}{n} \right) = PQ \left(\frac{N-n}{N-1} \right) \frac{1}{n}, \quad (2.1.18)$$

de aquí se puede deducir fácilmente la varianza de \hat{A} .

$$V(\hat{A}) = V(Np) = N^2 V(p) = N^2 PQ \left(\frac{N-n}{N-1} \right) \frac{1}{n}. \quad (2.1.19)$$

Una estimación insesgada de la varianza de p , derivada de la muestra es *Cochran (1984, pág 81)*:

$$\hat{V}(p) = \frac{N-n}{(n-1)N} pq. \quad (2.1.20)$$

El intervalo de confianza de p con **aproximación normal** es el siguiente *Cochran (1984, pág 88)*:

$$p \pm \left[t \sqrt{1-f} \sqrt{\frac{pq}{n-1}} + \frac{1}{2n} \right], \quad (2.1.21)$$

donde $f = n/N$ y t es el desvío normal correspondiente a la probabilidad de confianza. El término $1/2n$ es una corrección por continuidad que mejora la aproximación, el uso de $\sqrt{pq/n}$ raras veces produce diferencias apreciables, también existen otros métodos, el método presentado aquí es el de **aproximación normal**. Para tener idea si se tiene una distribución normal para las p posibles, se tiene una regla empírica: se tiene buena aproximación si $np > 5$ y $n(1-p) > 5$. En el caso del muestreo estratificado, la normalidad tiene más posibilidad porque se compensan los errores de un estrato a otro, además son más homogéneos los estratos. También existen tablas que dependen de n y p para saber cuando es aplicable la distribución normal sobre todo cuando p es extremadamente pequeña; asimismo existe el método de aproximaciones binomiales por ejemplo ver *Cochran (1984, pág 89)*.

2.1.6. MUESTREO INVERSO

El tamaño de muestra para valores entre 0.3 y 0.7 el valor anticipado de P no es trascendente, sin embargo para atributos raros, el tamaño de muestra se ve afectado seriamente por una pequeña variación, por ejemplo para población infinita donde $P = 10\%$ se tiene una $n_0 = 6,724$, mientras que para $P = 5\%$ se requiere una $n_0 = 14,195$, un crecimiento de 111.1%, esto obviamente implica mayores costos y esfuerzo. Para estas situaciones, estimar un valor pequeño de P , donde P esperado no es muy conocido, el método de *Haldane (1945)* del muestreo inverso tiene gran ventaja, consiste en un muestreo continuo hasta que m de los atributos raros se hayan encontrado en la muestra.

Si n es el tamaño de la muestra en el que aparece el m -ésimo atributo raro, ($m > 1$), una estimación insesgada de P es:

$$p = \frac{m-1}{n-1}. \quad (2.1.22)$$

Para N muy grande, P pequeña y $m \geq 10$, una aproximación de $V(p)$:

$$V(p) \cong \frac{mP^2Q}{(m-1)^2}, \quad (2.1.23)$$

el estimador de la varianza de p *Sukhatme et al (1984, pág 52)*:

$$\hat{V}(p) = \frac{p(1-p)}{n-2} \left(1 - \frac{n-1}{N} \right) \quad (2.1.24)$$

y

$$cv(p) \cong \frac{\hat{m}Q}{m-1} < \frac{\hat{m}}{m-1}. \quad (2.1.25)$$

Si P es pequeño el límite superior será lo suficientemente cercano al $cv(p)$. De este modo se puede controlar el valor del $cv(p)$ sin conocer anticipadamente P . Si $m = 47$, $cv(p) < 15\%$; sin embargo, para $cv(p) < 10\%$ se requieren de 102 casos. Con el método de muestreo inverso n es una variable aleatoria y definitivamente será grande si P es pequeña.

2.2 MUESTREO ALEATORIO ESTRATIFICADO

2.2.1. INTRODUCCION

En este tipo de muestreo, la población de N unidades se divide primero en subpoblaciones de N_1, N_2, \dots, N_L unidades, respectivamente. Las subpoblaciones tienen las siguientes características, no se traslapan los estratos y en su conjunto comprenden a toda la población, por lo tanto podemos deducir que:

$$N_1 + N_2 + N_3 + \dots + N_L = N.$$

Las subpoblaciones se denominan estratos. Cosa importante para que el muestreo estratificado nos presente la precisión requerida es que cada uno de los N_h sean conocidos. Un punto deseable de los estratos, es que dentro de cada uno de ellos, los individuos tengan características similares. Los tamaños de muestra dentro de cada uno de los estratos son n_1, n_2, \dots, n_L , respectivamente. Después de determinado el tamaño, se extrae la muestra de cada uno de los estratos, las extracciones deben de hacerse independientes en cada uno de los estratos.

Las principales razones para usar la técnica del muestreo aleatorio estratificado *Cochran (1984, pág 125)* son las siguientes:

- 1.- Si los datos deseados deben tener una precisión conocida en algunos dominios de estudio de la población, es decir, que ciertas subpoblaciones tengan representatividad por sí mismo; es aconsejable tratar cada dominio de estudio como una población por derecho propio.
- 2.- Por conveniencia administrativa, puede ser necesario el uso de la estratificación, así por ejemplo, la agencia que realiza una encuesta, podría tener sucursales en el campo, cada una de las cuales supervisaría la encuesta de una parte de la población.
- 3.- Los problemas de muestreo pueden tener marcadas diferencias en diversas partes de la población. Con poblaciones humanas, las personas que viven en instituciones (como hoteles, hospitales, cárceles) se colocan en un estrato distinto de los que viven en un lugar establecido a casas ordinarias.

4.- La estratificación nos puede ayudar a una ganancia en la precisión de las estimaciones de las características a medir de la población total. La idea es que si tenemos un universo heterogéneo, pero este tiene subpoblaciones, en la que cada subpoblación tiene internamente un comportamiento homogéneo, esto es lo que el nombre de estrato sugiere, con la división en capas. Si cada estrato es homogéneo, es decir, que la varianza dentro de cada estrato es pequeña o que el comportamiento que existe en cada individuo y otro la diferencia es mínima, una muestra pequeña puede dar una estimación precisa del estrato, así, posteriormente podrán agruparse las estimaciones de una estimación más precisa para toda la población.

2.2.2 NOTACION Y ESTIMADORES

El subíndice i representa la unidad dentro del estrato, h denota el estrato de los L que suponemos, los símbolos que a continuación se citan se refieren al estrato h , *Cochran (1984, págs 126-127)*.

$$N = \sum_{i=1}^L N_h \quad \text{La suma de las unidades por estrato nos da el total} \quad (2.2.1)$$

N_h = número total de unidades en el estrato h .

n_h = número de unidades en la muestra en el estrato h .

y_{hi} = valor obtenido de la i -ésima unidad del estrato h .

$W_h = N_h/N$ = Ponderador del estrato h .

$f_h = n_h/N_h$ = fracción de muestreo en el estrato h .

Media verdadera del estrato h

$$\bar{Y}_h = \frac{\sum_{i=1}^{N_h} y_{hi}}{N_h} \quad (2.2.2)$$

Media de muestra en el estrato h

$$\bar{y}_h = \frac{\sum_{i=1}^{n_h} y_{hi}}{n_h}, \quad (2.2.3)$$

cuadrados medios del estrato h

$$S_h^2 = \frac{1}{N_h - 1} \sum_{i=1}^{N_h} (y_{hi} - \bar{Y}_h)^2, \quad (2.2.4)$$

cuadrados medios en la muestra del estrato h

$$s_h^2 = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (y_{hi} - \bar{y}_h)^2. \quad (2.2.5)$$

2.2.3 ALGUNAS PROPIEDADES DE LOS ESTIMADORES

Después de obtener la media de población por unidad, la estimación insesgada usada en el muestreo estratificado es la siguiente *Cochran (1984, págs 127-130)*:

media cuando W_h es poblacional

$$\bar{y}_{st} = \frac{\sum_{h=1}^L N_h \bar{y}_h}{N} = \sum_{h=1}^L W_h \bar{y}_h, \quad (2.2.6)$$

media muestral

$$\bar{y} = \frac{\sum_{h=1}^L n_h \bar{y}_h}{n} = \sum_{h=1}^L w_h \bar{y}_h. \quad (2.2.7)$$

Es claro que ambas estimaciones son diferentes en términos generales, sin embargo son iguales cuando:

$$\frac{n_h}{n} = \frac{N_h}{N} \Leftrightarrow f_h = f.$$

Este tipo de muestreo se conoce como estratificación con asignación proporcional y da lugar a una muestra autoponderada, una muestra autoponderada es aquella en la que cada una de las unidades de muestreo (a las que se aplican las encuestas) tienen la misma probabilidad de ser

seleccionados. Este tipo de diseño es bondadoso en el sentido de que los estimadores de varianza son fáciles de calcular.

También es necesario citar la varianza para muestras estratificadas extraídas de manera independiente:

$$V(\bar{y}_{st}) = V\left(\sum_{h=1}^L W_h \bar{y}_h\right) = \sum_{h=1}^L W_h^2 V(\bar{y}_h). \quad (2.2.8)$$

Donde varianza de la media muestral y $f = n/N$ que se conoce como la fracción de muestreo es:

$$V(\bar{y}_h) = \frac{(N_h - n_h) S_h^2}{N_h n_h} = (1 - f) \frac{S_h^2}{n_h}. \quad (2.2.9)$$

La forma de lograr una estimación insesgada de la varianza de la media poblacional es sustituir cada S_h^2 por su respectivo s_h^2 *Sukhatme et al (1984, pág 107)*.

Para el caso del muestreo estratificado para proporciones se tiene:

$$V(p) = \frac{1}{N^2} \sum_{h=1}^L \frac{N_h^2 (N_h - n_h)}{N_h - 1} \frac{P_h Q_h}{n_h} \quad (2.2.10)$$

De esta manera el intervalo de confianza de la media queda:

$$p \pm t \cdot \sqrt{V(p)}, \quad (2.2.11)$$

y el intervalo de confianza para la media de algún estrato:

$$p_h \pm t \cdot \sqrt{V(p_h)}, \quad (2.2.12)$$

donde $V(p_h)$ es (2.1.18) pero sustituyendo cada valor por el de su respectivo estrato. Generalmente sólo se utiliza el intervalo de confianza de los dominios a estudiar, es decir, aquellos que tienen representatividad por las

necesidades del estudio; aquellos que no alcanzan representatividad, en la mayoría de los casos no tienen suficiente cantidad de encuestas para obtener resultados confiables por estrato y no son tomados en cuenta generalmente.

Existen a su vez diferentes tipos de asignación dado n para distribuir la cantidad de entrevistas en cada uno de los estratos, por ejemplo *Cochran (1984, pág 133-136)*.

i) Asignación proporcional

$$n_h = n \frac{N_h}{N}, \quad (2.2.13)$$

ii) Asignación óptima (considera los costos por entrevista por estrato)

$$n_h = n \frac{W_h S_h / \sqrt{c_h}}{\sum_{h=1}^L W_h S_h / \sqrt{c_h}}, \quad (2.2.14)$$

donde c_h es el costo por unidad de muestreo en el estrato h .

iii) Asignación de *Neyman (1934)*, *Tschuprow (1923)*

$$n_h = n \frac{W_h S_h}{\sum_{h=1}^L W_h S_h}. \quad (2.2.15)$$

2.3 MUESTREO POR CONGLOMERADOS

2.3.1. INTRODUCCION

Hay dos razones principales para la aplicación del muestreo por conglomerados *Cochran (1984, pág 289)*. Aunque la primera intención sea la de usar los elementos como unidades de muestreo, se ha encontrado que para muchas encuestas no se tiene una lista confiable de los elementos de la población, en el caso concreto de México existe el padrón electoral con un listado de las personas y sus datos particulares. En México no existe acceso a listas completas y actualizadas de la población, las viviendas o las granjas en grandes regiones geográficas. Sin embargo, a partir de los mapas de la región, se puede dividirla en unidades de área, como serían las manzanas de una ciudad y los terrenos con fronteras identificables rápidamente en las zonas rurales, en México se cuenta con la base de datos del Inegi que divide a las zonas urbanas en unidades llamadas Área Geoestadística Básica (Ageb) que son divisiones más pequeñas que el municipio y de fácil identificación, a las localidades con menos de 2,500 habitantes el Inegi las clasifica como rurales. También se cuenta con la lista de las Secciones Electorales con las que cuenta el país.

Aunque se cuente con una lista de elementos, suele ocurrir que por consideraciones económicas se utiliza una unidad conglomerada mayor, en el caso de México suelen usarse las Ageb's o las secciones electorales, para un tamaño de muestra dado, una unidad de muestreo pequeña suele ser más precisa que una unidad grande. Por ejemplo, una muestra aleatoria simple de 600 elementos cubre una ciudad más uniformemente que una muestra de 40 Ageb's con un promedio de 15 elementos por Ageb. Desafortunadamente los costos son mayores para localizar 600 elementos que localizar 40 Ageb's y visitar todas las casas de dicha Ageb (conglomerado). Al poner la vertiente que existe entre el costo y la precisión, la unidad mayor puede resultar más conveniente, en cuestiones de tiempo es más propicio usar un muestreo por conglomerados que buscar todos los elementos.

2.3.2 NOTACION Y ESTIMADORES

Si el universo U se divide en N conglomerados constituidos por M_i elementos, $i = 1, \dots, N$ y seleccionamos n de los N conglomerados por MAS entonces la notación a usar es *Sukhatme et al (1984, págs 290-292)*. Este autor presenta cuatro estimadores, aquí se muestran dos que son los más comunes en la práctica y son útiles para esta tesis, la ventaja que tiene el estimador de razón es que no es necesario conocer el tamaño de todos los conglomerados, el primer estimador presentado es insesgado y tiene gran utilidad en las encuestas pre y post-electorales.

Y_{ij} valor de la característica a investigar en el j -ésimo elemento del i -ésimo conglomerado, donde $j = 1, 2, \dots, M_i$; $i = 1, 2, \dots, N$.

M_i cantidad de elementos de segunda etapa en el i -ésimo elemento de primera etapa; $i = 1, 2, \dots, N$.

$\bar{M}_N = \frac{1}{N} \sum_{i=1}^N M_i$; tamaño promedio de unidades de segunda etapa.

$\bar{M}_n = \frac{1}{n} \sum_{i=1}^n M_i$, tamaño promedio de unidades de segunda etapa en muestra.

$$\bar{Y}_{i\cdot} = \frac{1}{M_i} \sum_{j=1}^{M_i} Y_{ij}, \text{ media del conglomerado } i. \quad (2.3.1)$$

$$\bar{\bar{Y}} = \frac{1}{N} \sum_{i=1}^N \bar{Y}_{i\cdot}, \text{ media de las medias que se obtienen en cada conglomerado.}$$

$$U_{ij} = \frac{M_i Y_{ij}}{M}, \quad \bar{U}_{i\cdot} = \frac{1}{M_i} \sum_{j=1}^{M_i} U_{ij} = \frac{M_i \bar{Y}_{i\cdot}}{M}.$$

De esta forma se obtiene:

$$\bar{\bar{Y}} = \frac{\sum_{i=1}^N M_i \bar{Y}_{i\cdot}}{\sum_{i=1}^N M_i} = \frac{1}{N} \sum_{i=1}^N \frac{M_i \bar{Y}_{i\cdot}}{M} = \frac{1}{N} \sum_{i=1}^N \bar{U}_{i\cdot} = \bar{\bar{U}} \quad (2.3.2)$$

i) estimador insesgado *Sukhatme et al (1984, pág 291)*:

$$\bar{y}_u = \bar{u} = \frac{\frac{1}{n} \sum_{i=1}^n M_i \bar{Y}_i}{\bar{M}_N} \quad (2.3.3)$$

La varianza del estimador queda determinada por:

$$V(\bar{y}_u) = \left(\frac{1}{n} - \frac{1}{N} \right) S_u^2, \quad (2.3.4)$$

donde:

$$S_u^2 = \frac{1}{N-1} \sum_{i=1}^N \left(\frac{M_i \bar{Y}_i}{\bar{M}} - \bar{Y} \right)^2, \quad (2.3.5)$$

y el estimador de (2.3.5) es (2.3.6):

$$s_u^2 = \frac{1}{n-1} \sum_{i=1}^n \left(\frac{M_i \bar{y}_i}{\bar{M}} - \bar{y} \right)^2, \quad (2.3.6)$$

de esta manera una estimación de la varianza se obtiene sustituyendo (2.3.5) por (2.3.6) en (2.3.4).

ii) estimador de razón *Sukhatme et al, (1984, pág 291)*:

$$\bar{y}_r = \frac{\bar{u}}{\bar{v}} = \frac{\sum_{i=1}^n M_i \bar{Y}_i}{\sum_{i=1}^n M_i}, \quad (2.3.7)$$

una aproximación de la varianza del estimador (2.3.7) es:

$$ECM(\bar{y}_r) \cong \left(\frac{1}{n} - \frac{1}{N} \right) S_r^2, \quad (2.3.8)$$

donde:

$$S_r^2 = \sum_{i=1}^N \frac{M_i^2}{M^2} \frac{(Y_{i\cdot} - \bar{Y})^2}{N-1}, \quad (2.3.9)$$

y el estimador de (2.3.9) es (2.3.10):

$$s_r^2 = \sum_{i=1}^n \frac{M_i^2}{M_n^2} \frac{(\bar{y}_{i\cdot} - \bar{y})^2}{n-1}. \quad (2.3.10)$$

De esta manera un estimador de la varianza se obtiene sustituyendo (2.3.10) por (2.3.9) en (2.3.8). Cabe mencionar que el promedio poblacional para el estimador de razón de la muestra no se conoce con anticipación.

No existen fórmulas exactas para la varianza de los estimadores de razón y el sesgo, sólo existen aproximaciones que son válidas para muestras grandes, el estimador de razón es consistente y entre mayor sea la muestra el sesgo tiende a despreciarse. Si la muestra es grande la razón es casi normalmente distribuida y se puede utilizar una aproximación de la varianza sin problema alguno. Asimismo, existe una cota para el sesgo del estimador de razón *Kish, Leslie (1965, pág. 249)*:

$$\text{sesgo}(\bar{y}_r) \leq \sqrt{\bar{V}(\bar{y}_r)} \sqrt{\frac{\bar{V}(y)}{Y}},$$

donde

$$V(y) = \left(\frac{1}{n} - \frac{1}{N} \right) \frac{1}{N-1} \sum_{i=1}^N \left(Y_i - \frac{Y}{N} \right)^2,$$

Y_i Total de la característica a investigar en el i -ésimo conglomerado, $i = 1, 2, \dots, N$.

Y Total de la característica a investigar de la población.

2.4 SUBMUESTREO O MUESTREO BIETAPICO

2.4.1 INTRODUCCION

En el capítulo anterior se presentó el muestreo por conglomerados, las unidades de muestreo son estos mismos, en el cual el conglomerado es censado. Cabe destacar que entre más grande sea el tamaño del conglomerado la eficacia del proceso decrece, siempre y cuando los elementos sean las unidades de muestreo, por esto, es mejor distribuir los elementos en varios conglomerados, que concentrados en uno solo.

Un submuestreo es aquel que tiene dos etapas de selección, la primera etapa son los conglomerados y la segunda los elementos dentro del conglomerado. Los conglomerados que constituyen las unidades de primera etapa se les llama unidades primarias y los elementos que constituyen la segunda etapa se les llama unidades secundarias o subunidades, desde luego que también existen muestreos de varias etapas, a dichos diseños se les llama multietápico. Un ejemplo de muestreo trietápico lo representa un estudio electoral del estado de Jalisco, donde las unidades primarias serían los 124 municipios, las unidades secundarias serían las secciones electorales dentro de cada uno de los municipios seleccionados y las unidades de tercera etapa serían los individuos dentro de cada sección electoral seleccionada.

2.4.2. NOTACION

En algunos casos las unidades primarias son de diferente tamaño, sin embargo, se da el caso en el que los conglomerados tienen el mismo tamaño, cada modelo tiene su propia notación. Por conveniencia aquí se presenta el de unidades de diferente tamaño *Sukhatme et al (1984, págs 323-325)* debido a que es un estimador común en la práctica.

M_i = Número de unidades secundarias en la i -ésima unidad primaria,
 $i= 1,2,\dots,N$

Y_{ij} = Valor de la característica a investigar en el j -ésimo elemento del i -ésimo conglomerado, donde $j = 1, \dots, M_i; i = 1, \dots, n$.

tamaño promedio de unidades de segunda etapa

$$\bar{M} = \frac{1}{N} \sum_{i=1}^N M_i, \quad (2.4.1)$$

la media de las observaciones en la i -ésima unidad primaria

$$\bar{Y}_{i\cdot} = \frac{1}{M_i} \sum_{j=1}^{M_i} Y_{ij}, \quad (2.4.2)$$

la media poblacional:

$$\bar{Y} = \bar{Y}_{\cdot\cdot} = \frac{1}{NM} \sum_{i=1}^N M_i \bar{Y}_{i\cdot}, \quad (2.4.3)$$

$$S_{iv}^2 = \frac{1}{M_i - 1} \sum_{j=1}^{M_i} (Y_{ij} - \bar{Y}_{i\cdot})^2, \quad (2.4.4)$$

$$S_{by}^2 = \frac{1}{N - 1} \sum_{i=1}^N \left(\frac{M_i}{M} \bar{Y}_{i\cdot} - \bar{Y} \right)^2, \quad (2.4.5)$$

$$U_{ij} = \frac{M_j}{M} Y_{ij}, \quad (2.4.6)$$

$$\bar{U}_{i\cdot} = \frac{1}{M_i} \sum_{j=1}^{M_i} U_{ij} = \frac{M_i}{M} \bar{Y}_{i\cdot}, \quad (2.4.7)$$

$$\bar{U} = \frac{1}{N} \sum_{i=1}^N \bar{U}_{i\cdot} = \bar{Y}_{\cdot\cdot}, \quad (2.4.8)$$

$$S_{iv}^2 = \frac{1}{M_i - 1} \sum_{j=1}^{M_i} (U_{ij} - \bar{U}_{i\cdot})^2 = \frac{M_i^2}{M^2} S_{iv}^2, \quad (2.4.9)$$

$$S_v^2 = \frac{1}{N-1} \sum_{i=1}^N (\bar{U}_i - \bar{U})^2 = S_{by}^2. \quad (2.4.10)$$

Sea m_i el número de unidades de segunda etapa elegidas de la i -ésima unidad primaria muestreada.

$$\bar{m} = \frac{1}{n} \sum_{i=1}^n m_i, \quad (2.4.11)$$

$$\bar{y}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} y_{ij}, \quad (2.4.12)$$

$$\bar{u}_i = \frac{1}{m_i} \sum_{j=1}^n u_{ij} = \frac{1}{m_i} \sum_{j=1}^n \frac{M_j}{M} y_{ij}, \quad (2.4.13)$$

$$\bar{u} = \frac{1}{n} \sum_{i=1}^n \bar{u}_i = \frac{1}{n} \sum_{i=1}^n \frac{M_i}{M} \bar{y}_i. \quad (2.4.14)$$

2.4.3 ESTIMADORES

Sukhatme et al (1984, págs 325-331) presenta cuatro estimadores, aquí se presentan tres de los cuales el segundo y el tercero son utilizados en esta tesis, el primero es sólo para mostrar alternativas a los utilizados:

(i) Primer estimador

$\hat{\bar{Y}}$ es un estimador sesgado de \bar{Y} ,

$$\hat{\bar{Y}} = \frac{1}{n} \sum_{i=1}^n \bar{y}_i. \quad (2.4.15)$$

La varianza del estimador es:

$$V(\hat{Y}) = \left(\frac{1}{n} - \frac{1}{N}\right) \frac{1}{N-1} \sum_{i=1}^N \left(\bar{Y}_{i\cdot} - \frac{1}{N} \sum_{i=1}^N \bar{Y}_{i\cdot}\right)^2 + \frac{1}{nN} \sum_{i=1}^N \left(\frac{1}{m_i} - \frac{1}{M_i}\right) S_{y_i}^2, \quad (2.4.16)$$

donde $S_{y_i}^2$ es (2.4.4) y para el ECM necesitamos el sesgo al cuadrado que es la diferencia de la media real y la estimada.

$$\text{sesgo} = \frac{1}{N} \sum_{i=1}^N \bar{Y}_{i\cdot} - \frac{1}{N} \sum_{i=1}^N N_i \bar{Y}_i,$$

y el estimador de (2.4.16) es:

$$\hat{V}(\hat{Y}) = \left(\frac{1}{n} - \frac{1}{N}\right) \frac{1}{n-1} \sum_{i=1}^n \left(\bar{y}_{i\cdot} - \frac{1}{n} \sum_{i=1}^n \bar{y}_{i\cdot}\right)^2 + \frac{1}{nN} \sum_{i=1}^n \left(\frac{1}{m_i} - \frac{1}{M_i}\right) s_{y_i}^2,$$

(ii) Segundo estimador:

\hat{Y}_{u2} es un estimador insesgado de \bar{Y}

$$\hat{Y}_{u2} = \frac{1}{n} \sum_{i=1}^n \frac{M_i}{M} \bar{y}_{i\cdot} = \frac{1}{n} \sum_{i=1}^n \bar{u}_i = \bar{u}, \quad (2.4.17)$$

la varianza del estimador es

$$\begin{aligned} V(\hat{Y}_{u2}) &= \left(\frac{1}{n} - \frac{1}{N}\right) S_u^2 + \frac{1}{nN} \sum_{i=1}^N \left(\frac{1}{m_i} - \frac{1}{M_i}\right) S_{u_i}^2 \\ &= \left(\frac{1}{n} - \frac{1}{N}\right) \frac{1}{(N-1)} \sum_{i=1}^N \left(\frac{M_i}{M} \bar{Y}_{i\cdot} - \bar{Y}\right)^2 + \frac{1}{nN} \sum_{i=1}^N \left(\frac{1}{m_i} - \frac{1}{M_i}\right) \left(\frac{M_i}{M}\right)^2 S_{y_i}^2, \end{aligned} \quad (2.4.18)$$

el estimador de la varianza del estimador es:

$$\hat{V}(\hat{Y}_{u2}) = \left(\frac{1}{n} - \frac{1}{N}\right) s_u^2 + \frac{1}{nN} \sum_{i=1}^n \left(\frac{1}{m_i} - \frac{1}{M_i}\right) s_{u_i}^2, \quad (2.4.19)$$

donde s_u^2 es (2.3.6) y

$$\begin{aligned} s_u^2 &= \frac{1}{m_i - 1} \sum_{j=1}^{m_i} (y_{ij} - \bar{y}_{i\cdot})^2, \\ s_{u_i}^2 &= \frac{M_i^2}{M^2} s_{y_i}^2. \end{aligned}$$

(iii) Tercer estimador

\hat{Y}_{r2} es un estimador sesgado de \bar{Y}

$$\hat{Y}_{r2} = \frac{\sum_{i=1}^n M_i \bar{y}_i}{\sum_{i=1}^n M_i}, \quad (2.4.20)$$

El ECM del estimador está dado en (2.4.21), donde S_w^2 es (2.4.4)

$$ECM(\hat{Y}_2) = \left(\frac{1}{n} - \frac{1}{N}\right) S_w^2 + \frac{1}{nN} \sum_{i=1}^n \left(\frac{M_i}{\bar{M}_n}\right)^2 \left(\frac{1}{m_i} - \frac{1}{M_i}\right) S_w^2, \quad (2.4.21)$$

una aproximación al estimador de varianza (ECM) se da por

$$E\hat{C}M(\hat{Y}_2) = \left(\frac{1}{n} - \frac{1}{N}\right) s_w^2 + \frac{1}{nN} \sum_{i=1}^n \left(\frac{M_i}{\bar{M}_n}\right)^2 \left(\frac{1}{m_i} - \frac{1}{M_i}\right) s_w^2, \quad (2.4.22)$$

donde:

$$s_w^2 = \frac{1}{n-1} \sum_{i=1}^n \left(\frac{M_i}{\bar{M}_n} \bar{y}_i - \bar{y}\right)^2, \quad (2.4.23)$$

donde \bar{u} es (2.4.14)

$$\bar{M}_n = \frac{1}{n} \sum_{i=1}^n M_i. \quad (2.4.24)$$

El sesgo del primer estimador es provocado porque las probabilidades de selección de las subunidades varían de una unidad a otra, es decir que las unidades primarias son de diferente tamaño. Si M_i tiene una variación considerable y además está correlacionada con la característica bajo estudio, el sesgo se puede depreciar, pero es muy importante que estas condiciones se cumplan antes de aceptar (2.4.15).

El segundo estimador, cumple con una bondad estadística deseable, no tiene sesgo, al ver la varianza, el primer sumando de (2.4.18) depende únicamente de la variación de los totales entre los conglomerados. Comparado con la varianza del estimador (2.4.15) se puede demostrar que (2.4.18) es mayor casi siempre que el estimador sesgado (2.4.16), siempre y cuando las condiciones de que la correlación entre el tamaño del conglomerado y su media sea positiva y que el sesgo del primer estimador (2.4.15) sea insignificante. De igual manera el segundo sumando de (2.4.18) será mayor al de (2.4.16) a condición de que exista correlación positiva entre M_i y S_i^2 ,. *Sukhatme et al (1984, págs 327-328)*.

El tercer estimador es una media ponderada de las medias de conglomerados, es un estimador de razón, se compone de la razón de dos variables aleatorias. Este estimador también es sesgado pero el sesgo se reduce conforme n crece, además este estimador es consistente, una ventaja es que no es necesario conocer la media del tamaño de las unidades de primera etapa para calcularlo.

Por lo general cuando el tamaño de los conglomerados varía considerablemente, quizá el estimador de razón es el más eficiente de los tres, siempre y cuando n sea grande y la variable bajo estudio y tenga un alto grado de correlación con la variable auxiliar x *Sukhatme et al (1984, pág 333)*.

El cuarto estimador no mencionado aquí es el de regresión, ver *Sukhatme et al (1984, págs 333-336)*.

2.5 SUBMUESTREO ESTRATIFICADO

2.5.1 INTRODUCCION

Es importante conocer la forma en que interactúan el muestreo estratificado y el submuestreo, en el capítulo anterior (2.4) se explicó el muestreo bietápico, se mostraron tres de cuatro estimadores que presenta *Sukhatme* y en el capítulo (2.2) se presentó el muestreo aleatorio estratificado. El objetivo de este capítulo es ver la forma en que interactúan estos dos diseños, la forma en que interactúan es primero estratificar el universo y después dentro de cada estrato se aplica el submuestreo correspondiente de tal manera que de las unidades primarias, se toma una muestra, seleccionada la unidad primaria se seleccionan de manera aleatoria las unidades secundarias de las unidades primarias seleccionadas. Un ejemplo de esto sería tomar un estudio de la Zona Metropolitana de Guadalajara, los estratos serían los municipios de Guadalajara, Tlaquepaque, Tonalá y Zapopan; las unidades primarias serían las secciones electorales y las unidades secundarias los individuos seleccionados que habitan dentro de la sección electoral seleccionada.

2.5.2 NOTACION

Si la población la dividimos en L estratos *Sukhatme et al (1984, págs 352-354)*, sea:

N_t = número de unidades primarias en el t -ésimo estrato de la población
 $t = 1, \dots, L$.

$$N = \sum_{t=1}^L N_t, \quad (2.5.1)$$

M_{it} = número de unidades en la segunda etapa de la i -ésima unidad primaria del t -ésimo estrato $t = 1, \dots, L, i = 1, \dots, N_t$.

Y_{ij} = valor del atributo en la j -ésima unidad secundaria de la i -ésima unidad primaria dentro del t -ésimo estrato $t = 1, \dots, L, i = 1, \dots, N_t, j = 1, \dots, M_{it}$.

$$\bar{Y}_{i\cdot} = \frac{1}{M_{i\cdot}} \sum_{j=1}^{M_{i\cdot}} Y_{ij}, \quad (2.5.2)$$

$$\bar{Y}_{i\cdot\cdot} = \frac{1}{M_{i0}} \sum_{j=1}^{N_i} \sum_{k=1}^{M_{ij}} Y_{ijk} = \frac{1}{N_i M_{i\cdot}} \sum_{j=1}^{N_i} \sum_{k=1}^{M_{ij}} Y_{ijk}, \quad (2.5.3)$$

$$\bar{Y} = \bar{Y}_{\cdot\cdot\cdot} = \frac{\sum_{i=1}^L M_{i0} \bar{Y}_{i\cdot\cdot}}{\sum_{i=1}^L M_{i0}}, \quad (2.5.4)$$

$$\lambda_i = \frac{M_{i0}}{\sum_{i=1}^L M_{i0}} = \frac{M_{i0}}{M_0}, \quad (2.5.5)$$

n_t = número de unidades primarias en muestra del t -ésimo estrato

$$n = \sum_{t=1}^L n_t, \quad (2.5.6)$$

m_{it} = número de unidades de la segunda etapa en muestra de la i -ésima unidad primaria del t -ésimo estrato

$$S_{iy}^2 = \frac{1}{(M_{i\cdot} - 1)} \sum_{j=1}^{M_{i\cdot}} (Y_{ij} - \bar{Y}_{i\cdot})^2, \quad (2.5.7)$$

$$S_{i0}^2 = \frac{1}{(N_i - 1)} \sum_{j=1}^{N_i} \left(\frac{M_{ij}}{M_{i\cdot}} \bar{Y}_{i\cdot\cdot} - \bar{Y}_{i\cdot\cdot} \right)^2. \quad (2.5.8)$$

2.5.3 ESTIMADORES

El estimador (2.5.11) que se presenta a continuación es un estimador insesgado de la media poblacional.

$$\bar{y}_i = \frac{1}{m_{it}} \sum_{j=1}^{m_{it}} y_{ij}, \quad (2.5.9)$$

$$\bar{y}_i = \frac{1}{M_i n_i} \sum_{j=1}^{n_i} \frac{M_{ij}}{m_{ij}} \sum_{j=1}^{m_{ij}} y_{ij} = \frac{1}{n_i} \sum_{j=1}^{n_i} \frac{M_{ij}}{M_i} \bar{y}_{ij}, \quad (2.5.10)$$

$$\bar{y}_{w2} = \sum_{i=1}^L \lambda_i \bar{y}_{is}, \quad (2.5.11)$$

además:

$$s_{iy}^2 = \frac{1}{(m_{ii} - 1)} \sum_{j=1}^{m_{ii}} (y_{ij} - \bar{y}_{ij})^2, \quad (2.5.12)$$

$$s_{ib}^2 = \frac{1}{(n_i - 1)} \sum_{j=1}^{n_i} \left(\frac{M_{ij}}{M_i} \bar{y}_{ij} - \bar{y}_i \right)^2, \quad (2.5.13)$$

a continuación se presenta la varianza del estimador \bar{Y} .

$$V(\bar{y}_{w2}) = \sum_{i=1}^L \lambda_i^2 \left[\left(\frac{1}{n_i} - \frac{1}{N_i} \right) S_{ib}^2 + \frac{1}{n_i N_i} \sum_{j=1}^{n_i} \left(\frac{M_{ij}}{M_i} \right)^2 \left(\frac{1}{m_{ij}} - \frac{1}{M_{ij}} \right) S_{iy}^2 \right], \quad (2.5.14)$$

un estimador insesgado de la varianza:

$$\hat{V}(\bar{y}_{w2}) = \sum_{i=1}^L \lambda_i^2 \left[\left(\frac{1}{n_i} - \frac{1}{N_i} \right) s_{ib}^2 + \frac{1}{n_i N_i} \sum_{j=1}^{n_i} \left(\frac{M_{ij}}{M_i} \right)^2 \left(\frac{1}{m_{ij}} - \frac{1}{M_{ij}} \right) s_{iy}^2 \right]. \quad (2.5.15)$$

La expresión (2.5.14) se puede escribir en forma resumida utilizando (2.4.18) por cada estrato

$$V(\bar{y}_{w2}) = \sum_{i=1}^L \lambda_i^2 V(\bar{y}_{is}), \quad (2.5.16)$$

y para el estimador de la varianza del estimador (2.5.16).

$$\hat{v}(\bar{y}_{w2}) = \sum_{i=1}^L \lambda_i^2 \hat{v}(\bar{y}_{is}), \quad (2.5.17)$$

donde

$$\hat{v}(\bar{y}_{is}) = \left(\frac{1}{n_i} - \frac{1}{N_i} \right) s_{ib}^2 + \frac{1}{N_i n_i} \sum_{j=1}^{n_i} \left(\frac{1}{m_{ij}} - \frac{1}{M_{ij}} \right) \left(\frac{M_{ij}}{M_i} \right)^2 s_{iy}^2. \quad (2.5.18)$$

Es claro que (2.5.15) = (2.5.17).

2.6. MUESTREO TRIETAPICO

2.6.1. INTRODUCCION

El muestreo trietápico es una práctica común en la actualidad en el diseño de las encuestas electorales, la ventaja más importante sobre el bietápico son los costos, un ejemplo sería un estado que tiene 9 municipios, si se desea hacer una encuesta electoral se está en posibilidades de visitar los 9 municipios, sin embargo, si el estado tiene 100 municipios el costo del operativo sería costoso si se visitaran todos, pero, de aprobarse éste sería un operativo en campo muy inadecuado por cuestiones de tiempo y traslados largos, este ejemplo es viable para aplicar un muestreo trietápico. Un ejemplo lo podemos suponer en una encuesta del estado de Jalisco, donde las unidades de primera etapa serían los municipios, las unidades de segunda etapa serían las secciones electorales y las unidades de última y tercera etapa sería los individuos seleccionados para la entrevista.

2.6.2. NOTACION

La notación y los estimadores de un muestreo de tres etapas con unidades de primera y segunda etapas de tamaños desiguales, se presentan a continuación: *Sukhatme et al (1984, págs 346-351)*

N = número total de unidades de primera etapa,

M_i = número de unidades de segunda etapa en la i -ésima unidad de primera etapa, $i = 1, \dots, N$.

B_{ij} = número de unidades de tercera etapa en la j -ésima unidad de segunda etapa de la i -ésima unidad de primera etapa, $j = 1, \dots, M_i$; $i = 1, \dots, N$.

Y_{ijk} = valor de la k -ésima unidad de tercera etapa en la j -ésima unidad de segunda etapa de la i -ésima unidad de primera etapa, $i = 1, \dots, N$, $j = 1, \dots, M_i$; $k = 1, \dots, B_{ij}$.

$$\bar{Y}_{ij\cdot} = \frac{1}{B_{ij}} \sum_{k=1}^{B_{ij}} Y_{ijk}, \quad (2.6.1)$$

$$\bar{B}_{i^*} = \frac{1}{M_i} \sum_{j=1}^{M_i} B_{ij}, \quad (2.6.2)$$

$$Q_i = M_i \bar{B}_{i^*}, \quad (2.6.3)$$

$$\sum_{i=1}^N Q_i = N \bar{Q}, \quad (2.6.4)$$

$$w_i = \frac{Q_i}{\bar{Q}}, \quad (2.6.5)$$

$$U_{ijk} = \frac{B_{ij}}{B_{i^*}} Y_{ijk}, \quad (2.6.6)$$

$$\bar{U}_{ij^*} = \frac{1}{B_{ij}} \sum_{k=1}^{B_{ij}} U_{ijk} = \frac{1}{B_{ij}} \sum_{k=1}^{B_{ij}} \frac{B_{ij}}{B_{i^*}} Y_{ijk} = \frac{B_{ij} \bar{Y}_{ij^*}}{B_{i^*}}, \quad (2.6.7)$$

$$\bar{Y}_{i^{**}} = \frac{\sum_{j=1}^{M_i} \sum_{k=1}^{B_{ij}} Y_{ijk}}{\sum_{j=1}^{M_i} B_{ij}} = \frac{1}{M_i} \sum_{j=1}^{M_i} \frac{B_{ij} \bar{Y}_{ij^*}}{B_{i^*}} = \frac{1}{M_i} \sum_{j=1}^{M_i} \bar{U}_{ij^*}, \quad (2.6.8)$$

se define:

$$\bar{U}_i = \frac{1}{M_i} \sum_{j=1}^{M_i} \bar{U}_{ij^*} = \bar{Y}_{i^{**}}, \quad (2.6.9)$$

$$\begin{aligned} \bar{Y} = \bar{Y}_{**} &= \frac{\sum_{i=1}^N \sum_{j=1}^{M_i} \sum_{k=1}^{B_{ij}} Y_{ijk}}{\sum_{i=1}^N \sum_{j=1}^{M_i} B_{ij}} = \frac{\sum_{i=1}^N \sum_{j=1}^{M_i} B_{ij} \bar{Y}_{ij^*}}{\sum_{i=1}^N M_i \bar{B}_i} \\ &= \frac{\sum_{i=1}^N M_i \bar{B}_{i^*} \bar{Y}_{i^{**}}}{N \bar{Q}} = \frac{1}{N} \sum_{i=1}^N \frac{Q_i \bar{Y}_{i^{**}}}{\bar{Q}}, \end{aligned} \quad (2.6.10)$$

de esta manera tenemos la media poblacional:

$$\begin{aligned}\bar{Y} &= \frac{\sum_{i=1}^N \sum_{j=1}^{M_i} B_{ij} \bar{Y}_{ij}}{\sum_{i=1}^N M_i \bar{B}_i} = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^{M_i} \frac{\bar{B}_i}{\bar{Q}} \bar{U}_{ij} \\ &= \frac{1}{N} \sum_{i=1}^N \frac{Q_i}{\bar{Q}} \bar{U}_i = \frac{1}{N} \sum_{i=1}^N w_i \bar{U}_i.\end{aligned}\quad (2.6.11)$$

Supongamos que:

n = número de unidades primarias seleccionadas en la muestra

m_i = número de unidades secundarias en la i -ésima unidad de primera etapa.

b_{ij} = número de unidades de tercera etapa seleccionadas de la j -ésima unidad de segunda etapa, de la i -ésima unidad de primera etapa.

$$\bar{u}_{ij} = \frac{1}{b_{ij}} \sum_{k=1}^{b_{ij}} u_{ijk} = \frac{1}{b_{ij}} \sum_{k=1}^{b_{ij}} \frac{B_{ijk}}{\bar{B}_i} y_{ijk} = \frac{B_{ij}}{\bar{B}_i} y_{ij}, \quad (2.6.12)$$

$$\bar{u}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} \bar{u}_{ij} = \frac{1}{m_i} \sum_{j=1}^{m_i} \frac{B_{ij}}{\bar{B}_i} y_{ij}, \quad (2.6.13)$$

$$\bar{u} = \frac{1}{n} \sum_{i=1}^n \frac{Q_i}{\bar{Q}} \bar{u}_i = \frac{1}{n} \sum_{i=1}^n w_i \bar{u}_i. \quad (2.6.14)$$

2.6.2. ESTIMADORES

Se presentan dos estimadores, el primero es insesgado y el segundo es de razón, estos estimadores se utilizaron en las encuestas que se presentan en esta tesis en el estado de Nuevo León, de allí su importancia de mostrarlos:

i) Estimador insesgado

$$\hat{\bar{Y}}_{u3} = \frac{1}{n} \sum_{i=1}^n w_i \frac{1}{m_i} \sum_{j=1}^{m_i} \frac{1}{b_{ij}} \sum_{k=1}^{b_{ij}} u_{ijk}, \quad (2.6.15)$$

si comparamos (2.6.15) con (2.6.16) es fácil ver que es insesgado

$$\bar{Y} = \frac{1}{N} \sum_{i=1}^N \frac{Q_i}{\bar{Q}} \frac{1}{M_i} \sum_{j=1}^{M_i} \frac{1}{B_{ij}} \sum_{k=1}^{B_{ij}} U_{ijk}. \quad (2.6.16)$$

La varianza queda determinada por:

$$V(\hat{Y}_{\bullet\bullet}) = \left(\frac{1}{n} - \frac{1}{N}\right) S_{b_i}^2 + \frac{1}{nN} \sum_{i=1}^N w_i^2 \left(\frac{1}{m_i} - \frac{1}{M_i}\right) S_{y_i}^2 + \frac{1}{nN} \sum_{i=1}^N \frac{w_i^2}{m_i M_i} \sum_{j=1}^{M_i} \frac{B_{y_j}^2}{\bar{B}_i^2} \left(\frac{1}{b_{y_j}} - \frac{1}{B_y}\right) S_{y_{y_j}}^2 \quad (2.6.17)$$

donde

$$S_{b_i}^2 = \frac{1}{(N-1)} \sum_{i=1}^N \left(\frac{Q_i}{Q} \bar{Y}_{i\bullet\bullet} - \bar{Y}\right)^2, \quad (2.6.18)$$

$$S_{y_i}^2 = \frac{1}{(M_i-1)} \sum_{j=1}^{M_i} \left(\frac{B_{y_j}}{\bar{B}_i} \bar{Y}_{y_j\bullet} - \bar{Y}_{i\bullet\bullet}\right)^2, \quad (2.6.19)$$

$$S_{y_{y_j}}^2 = \frac{1}{B_{y_j}-1} \sum_{k=1}^{B_{y_j}} \left(Y_{y_j k} - \bar{Y}_{y_j\bullet}\right)^2. \quad (2.6.20)$$

y el estimador de la varianza del estimador es:

$$\hat{V}(\hat{Y}_{\bullet\bullet}) = \left(\frac{1}{n} - \frac{1}{N}\right) s_{b_i}^2 + \frac{1}{nN} \sum_{i=1}^n w_i^2 \left(\frac{1}{m_i} - \frac{1}{M_i}\right) s_{y_i}^2 + \frac{1}{nN} \sum_{i=1}^n \frac{w_i^2}{m_i M_i} \sum_{j=1}^{m_i} \frac{B_{y_j}^2}{\bar{B}_i^2} \left(\frac{1}{b_{y_j}} - \frac{1}{B_y}\right) s_{y_{y_j}}^2, \quad (2.6.21)$$

donde:

$$s_{b_i}^2 = \frac{1}{(n-1)} \sum_{i=1}^n \left(\frac{Q_i}{Q} \bar{y}_{i\bullet\bullet} - \bar{y}\right)^2, \quad (2.6.22)$$

$$s_{y_i}^2 = \frac{1}{(m_i-1)} \sum_{j=1}^{m_i} \left(\frac{B_{y_j}}{\bar{B}_i} \bar{y}_{y_j\bullet} - \bar{y}_{i\bullet\bullet}\right)^2, \quad (2.6.23)$$

$$s_{y_{y_j}}^2 = \frac{1}{b_{y_j}-1} \sum_{k=1}^{b_{y_j}} \left(y_{y_j k} - \bar{y}_{y_j\bullet}\right)^2. \quad (2.6.24)$$

ii) Estimador de razón:

$$\hat{Y}_{r3} = (\hat{Y}_{u3} / \hat{X}_{u3}) \bar{X}, \quad (2.6.24)$$

donde \hat{Y}_{u3} es (2.6.15) y \hat{X}_{u3} es (2.6.15) pero sustituyendo y_{ijk} por x_{ijk} de esta manera x_{ijk} es una variable auxiliar de y_{ijk} , y para proporción $x_{ijk} = 1 \forall y_{ijk}$ si suponemos que $x_{ijk} = 1$ para todo i, j, k . entonces el estimador nos queda de la siguiente manera:

$$\bar{X} = \bar{X} \dots = \frac{\sum_{i=1}^N \sum_{j=1}^{M_i} \sum_{k=1}^{B_{ij}} X_{ijk}}{\sum_{i=1}^N \sum_{j=1}^{M_i} B_{ij}} = \frac{\sum_{i=1}^N \sum_{j=1}^{M_i} B_{ij}}{\sum_{i=1}^N \sum_{j=1}^{M_i} B_{ij}} = 1, \quad (2.6.25)$$

$$\hat{X}_{u3} = \frac{1}{n} \sum_{i=1}^n w_i \frac{1}{m_i} \sum_{j=1}^{m_i} \frac{1}{b_{ij}} \sum_{k=1}^{b_{ij}} \frac{B_{ij}}{B_{i*}} x_{ijk}, \quad (2.6.26)$$

y

$$\begin{aligned} \hat{Y}_{u3} &= \frac{1}{n} \sum_{i=1}^n w_i \frac{1}{m_i} \sum_{j=1}^{m_i} \frac{1}{b_{ij}} \sum_{k=1}^{b_{ij}} \frac{B_{ij}}{B_{i*}} y_{ijk} = \frac{1}{n} \sum_{i=1}^n w_i \frac{1}{m_i} \sum_{j=1}^{m_i} \frac{B_{ij}}{B_{i*}} \frac{\sum_{k=1}^{b_{ij}} y_{ijk}}{b_{ij}} \\ &= \frac{1}{n} \sum_{i=1}^n w_i \frac{1}{m_i} \sum_{j=1}^{m_i} \frac{B_{ij}}{B_{i*}} \bar{y}_{ij*}, \end{aligned} \quad (2.6.27)$$

de igual manera con:

$$\hat{X}_{u3} = \frac{1}{n} \sum_{i=1}^n w_i \frac{1}{m_i} \sum_{j=1}^{m_i} \frac{B_{ij}}{B_{i*}} x_{ij*} = \frac{1}{n} \sum_{i=1}^n w_i \frac{1}{m_i} \sum_{j=1}^{m_i} \frac{B_{ij}}{B_{i*}} = 1, \quad (2.6.28)$$

Entonces dividiendo (2.6.27) y (2.6.28) obtenemos:

$$\bar{y}_{r3} = \frac{\frac{1}{n} \sum_{i=1}^n w_i \frac{1}{m_i} \sum_{j=1}^{m_i} \frac{B_{ij}}{B_{i*}} \bar{y}_{ij*}}{\frac{1}{n} \sum_{i=1}^n w_i \frac{1}{m_i} \sum_{j=1}^{m_i} \frac{B_{ij}}{B_{i*}}} \quad (2.6.29)$$

Una aproximación del sesgo del estimador:

$$B(\hat{y}_{r3}) \cong \left(\frac{1}{n} - \frac{1}{N} \right) \left(\frac{S_{bx}^2}{\bar{X}^2} - \frac{S_{by}^2}{\bar{X}\bar{Y}} \right) + \frac{1}{nN} \sum_{i=1}^N w_i^2 \left(\frac{1}{m_i} - \frac{1}{M_i} \right) \left(\frac{S_{ix}^2}{\bar{X}^2} - \frac{S_{iy}^2}{\bar{X}\bar{Y}} \right) \\ + \frac{1}{nN} \sum_{i=1}^N \frac{w_i^2}{m_i M_i} \sum_{j=1}^{M_i} \left(\frac{B_{y_j}}{\bar{B}_{i^*}} \right)^2 \left(\frac{1}{b_{y_j}} - \frac{1}{B_{y_j}} \right) \left(\frac{S_{y_j^*}^2}{\bar{X}^2} - \frac{S_{y_j^*}}{\bar{X}\bar{Y}} \right), \quad (2.6.30)$$

la aproximación del error cuadrático medio para calcular la varianza nos queda determinado por:

$$ECM(\hat{y}_{r3}) \cong \left(\frac{1}{n} - \frac{1}{N} \right) (S_{by}^2 - 2RS_{bxy} + R^2 S_{bx}^2) + \frac{1}{nN} \sum_{i=1}^N w_i^2 \left(\frac{1}{m_i} - \frac{1}{M_i} \right) (S_{iy}^2 - 2RS_{ixy} + R^2 S_{ix}^2) \\ + \frac{1}{nN} \sum_{i=1}^N \frac{w_i^2}{m_i M_i} \sum_{j=1}^{M_i} \left(\frac{B_{y_j}}{\bar{B}_{i^*}} \right)^2 \left(\frac{1}{b_{y_j}} - \frac{1}{B_{y_j}} \right) (S_{y_j^*}^2 - 2RS_{y_j^*xy} + R^2 S_{y_j^*x}^2), \quad (2.6.31)$$

donde

$$S_{bx} = \frac{1}{(N-1)} \sum_{i=1}^N \left(\frac{Q_i}{\bar{Q}} \bar{X}_{i^*} - \bar{X} \right) \left(\frac{Q_i}{\bar{Q}} \bar{Y}_{i^*} - \bar{Y} \right), \quad (2.6.32)$$

$$S_{iy} = \frac{1}{M_i - 1} \sum_{j=1}^{M_i} \left(\frac{B_{y_j}}{\bar{B}_{i^*}} \bar{X}_{y_j^*} - \bar{X}_{i^*} \right) \left(\frac{B_{y_j}}{\bar{B}_{i^*}} \bar{Y}_{y_j^*} - \bar{Y}_{i^*} \right), \quad (2.6.33)$$

$$S_{y_j^*xy} = \frac{1}{(B_{y_j} - 1)} \sum_{k=1}^{B_{y_j}} (X_{y_jk} - \bar{X}_{y_j^*}) (Y_{y_jk} - \bar{Y}_{y_j^*}), \quad (2.6.34)$$

y S_{bx}^2 , S_{ix}^2 y S_{ijx}^2 son definidas en (2.6.32), (2.6.33), (2.6.34) pero sustituyendo las y 's por las x 's.

Los estimadores de S_{bxy} , S_{ixy} y S_{ijxy} son s_{bxy} , s_{ixy} y s_{ijxy} :

$$s_{y_j^*xy} = \frac{1}{b_{y_j} - 1} \sum_{k=1}^{b_{y_j}} (x_{y_jk} - \bar{x}_{y_j^*}) (y_{y_jk} - \bar{y}_{y_j^*}), \quad (2.6.35)$$

$$s_{ixy} = \frac{1}{(m_i - 1)} \sum_{j=1}^{m_i} \left(\frac{B_{y_j}}{\bar{B}_{i^*}} \bar{x}_{y_j} - \bar{x}_{i^*} \right) \left(\frac{B_{y_j}}{\bar{B}_{i^*}} \bar{y}_{y_j} - \bar{y}_{i^*} \right), \quad (2.6.36)$$

$$s_{bxy} = \frac{1}{n-1} \sum_{i=1}^n (w_i \bar{x}_i - \bar{x}_{w3}) (w_i \bar{y}_i - \bar{y}_{w3}), \quad (2.6.37)$$

un estimador consistente del ECM para calcular la varianza del estimador es:

$$\begin{aligned}
 \hat{V}(y_{r3}) = & \left(\frac{1}{n} - \frac{1}{N} \right) (s_{by}^2 - 2\hat{R}s_{by} + \hat{R}^2 s_{br}^2) \\
 & + \frac{1}{nN} \sum_{i=1}^n w_i^2 \left(\frac{1}{m_i} - \frac{1}{M_i} \right) (s_{iy}^2 - 2\hat{R}s_{iy} + \hat{R}^2 s_{ir}^2) \\
 & + \frac{1}{nN} \sum_{i=1}^n \frac{w_i^2}{m_i M_i} \sum_{j=1}^{m_i} \frac{B_{ij}^2}{\bar{B}_i^2} \left(\frac{1}{b_{ij}} - \frac{1}{B_{ij}} \right) (s_{ijy}^2 - 2\hat{R}s_{ijy} + \hat{R}^2 s_{ijr}^2)
 \end{aligned} \tag{2.6.38}$$

donde

$$\hat{R} = \frac{\bar{y}_{u3}}{\bar{x}_{u3}}$$

La razón de presentar estos estimadores es porque son los más usados en la práctica, al menos en mi experiencia en el Centro de Estudios de Opinión de la Universidad de Guadalajara, básicamente son tres estimadores de diferentes etapas combinado con el estratificado, el estimador de razón por ejemplo (2.4.20), es ampliamente utilizado en los conteos rápidos, el estimador insesgado por ejemplo (2.4.17) es utilizado en las encuestas preelectorales y las post-electtorales, la diferencia está en que en las encuestas preelectorales y post-electtorales conocemos la lista nominal de los puntos de muestreo (secciones electorales) y en el conteo rápido sabemos cuantos electores acudieron a votar después de haber terminado la jornada electoral, de allí las ventajas de utilizar uno u otro en la práctica. El estimador sesgado por ejemplo (2.4.15) que es una media no ponderada de medias, se utiliza cuando no se tiene algún marco muestral confiable, éste es utilizado en su mayoría en estudios de mercado.

2.7 EFECTO DE DISEÑO

2.7.1 INTRODUCCION

El efecto de diseño, es la razón de la varianza de un estimador bajo el esquema en que fue diseñado el muestreo y la varianza de un muestreo aleatorio simple de la misma muestra, es decir, los mismos elementos bajo distintos esquemas de muestreo; también se puede definir como la eficiencia relativa de un m.a.s. sobre un muestreo dado.

2.7.2 NOTACION

Leslie Kish (1989) propone dos definiciones:

$$Deff(\hat{\theta}) = \frac{V(\hat{\theta})}{V(\hat{\theta}_{mascr})}, \quad (2.7.1)$$

$$Deft^2(\hat{\theta}) = \frac{V(\hat{\theta})}{V(\hat{\theta}_{mascr})}, \quad (2.7.2)$$

donde

$Deff$ = Efecto de diseño donde se compara la eficiencia relativa de un m.a.s. sin reemplazo con el muestreo diseñado.

$Deft^2$ = Efecto de diseño donde se compara la eficiencia relativa de un m.a.s. con reemplazo con el muestreo diseñado.

$V(\hat{\theta})$ = varianza del estimador bajo el muestreo diseñado

El problema aquí, se da cuando no se conocen las varianzas de los estimadores, para el caso de la varianza del estimador bajo el muestreo diseñado, se puede estimar a partir de los datos muestrales, pero para estimar la varianza bajo un muestreo aleatorio simple, ésta debe expresarse en

términos de la varianza estimada bajo el diseño que se empleó, lo cual nos dificulta nuestra estimación.

Puede suceder que el estimador de θ sea sesgado pero consistente, si el tamaño de muestra es lo suficientemente grande el sesgo es negligible y el efecto del diseño se calcula con (2.7.1) ó (2.7.2) según sea el caso.

El efecto de diseño se utiliza básicamente para:

Calcular el tamaño de muestra. De tenerse el *Deff* de un muestreo complejo que se desea aplicar y el tamaño de muestra n requerido para un muestreo aleatorio simple, el tamaño de muestra requerido es:

$$n = n_{mas} Deff \quad (2.7.3)$$

Ahorrar cálculos. El efecto de diseño es una medida para efectos grandes y en la mayoría de los casos presenta la información para la población total, algunas encuestas requieren de información para varias variables y subpoblaciones, para no presentar varianza o error para cada caso se podría presentar un efecto de diseño general. Debe tenerse precaución en que los efectos no varíen mucho entre subpoblaciones o variables.

Comparar eficiencia de esquemas de muestreo complejos. Esto obviamente nos ayuda para decidir que modelo muestral utilizar y obtener la mayor precisión posible.

Encuestas periódicas. El efecto de diseño nos permite evaluar periodo a periodo cada modelo muestral utilizado en cada encuesta, este nos puede ayudar en obtener mayor precisión. De utilizarse el mismo modelo muestral encuesta a encuesta puede ajustarse la cantidad de encuestas periodo a periodo.

2.8 MUESTREO POR CUOTAS

El muestreo por cuotas fue desarrollado por *Cherington, Roper, Gallup y Crossley*, sus estudios de opinión fueron difundidos ampliamente después de la elección presidencial de Estados Unidos de 1936 *Cochran (1984)*. Este muestreo tiene características simples que de inmediato fue aceptado, incluirlo en una encuesta fue la moda, encuestas famosas lo habían utilizado, otra ventaja es que ofrece una disminución de costos. En los últimos años ha sido reemplazado por otras técnicas de muestreo probabilístico que han dado un importante avance en la materia. El muestreo por cuotas fue novedoso en la década de los cincuenta, que se debió a los costos bajos con los que operaba, éste es uno de los aspectos fundamentales en el uso del muestreo *Stephan et al (1958)*.

El 9 de Abril de 1992 *John Major* en Inglaterra ganó las elecciones y engañó a la industria encuestadora, en ese año encuestas por cuotas publicadas daban a la cabeza al partido laborista, con un punto porcentual de diferencia, pero al hacer el conteo de la votación el partido conservador obtuvo la victoria por una diferencia de 8 puntos porcentuales, de esta manera en 1992 fue el peor desastre de los últimos 50 años de la agencias encuestadoras *Ivor Crewe (1997)*. Desde entonces, se realizaron debates sobre las metodologías, entre las ventajas como se comentó del muestreo por cuota es el costo, pero de qué sirven costos bajos si los resultados no son acordes a la realidad, desde luego que se hicieron la pregunta entre los debates si era necesario cambiar al muestreo probabilístico y esto debería ocurrir rápidamente, *John Curtis & Nick Sparrow (1997)*, este tipo de desaciertos dejan en entredicho la credibilidad de los órganos competentes.

The Market Research Society por conducto de un panel de académicos y casas encuestadoras, identificaron 3 causas principales del error, *Ivor Crewe (1997)*:

- La primera fue un cambio repentino para los conservadores en la víspera, asimismo en el día de las elecciones internas el voto fue volátil, porque casi un 21% del electorado cambió de opinión durante la campaña, comparado con 19, 15 y 13% con las tres previas elecciones. Las entrevistas mostraron que en 36 horas acercándose el día de la elección algunos que habían declarado abstenerse de ir a votar, votaron y viceversa.

- La segunda fue un error de diseño de muestreo, los encuestadores fueron capacitados para entrevistar a determinado número de personas de acuerdo a su edad, género, clase social y estado laboral. Por lo general género y edad sólo se relacionan débilmente con el voto, pero la clase social corrobora y fortalece la información, sin embargo, es difícil para los encuestadores aplicar con precisión las encuestas ya que tendían a no encontrar *respondientes* de la clase social que se les requería.
- La tercera causa fue la espiral de silencio, un término conocido por la investigadora *Elizabeth Noelle Newman*, para la tendencia de la gente a evadir o permanecer en un clima hostil de opinión a sus alrededores, especialmente evitando responder la pregunta ¿Por quien va a votar?, respondiendo no sé o simplemente no contestaron.

Cabe destacar que las fuentes para definir las cuotas se obtenían de diferentes destinos y éstas no coincidían en las empresas encuestadoras, por ejemplo, *Gallup* utilizó cuotas con referencia a proyecciones de la oficina nacional de estadística. Las cuotas más comunes que aplican *MORI, Gallup, NOP, BSA, BEPS e ICM* fueron sexo, edad, ocupación, nivel socioeconómico y clase social, la suposición del uso de las cuotas es bajo el supuesto de lograr mayor dispersión de las encuestas *Ivor Crewe (1997)*. Una variable de cuota no necesariamente está relacionada con la característica de la entrevista, entonces para que definirla?, algunas empresas desean demasiada información detallada, esto ocasiona que se pierda el objetivo principal y además genera problemas en campo por la dificultad de completar la cuota, por consiguiente, deben hacerse solamente variables asociadas fuertemente con la característica de la entrevista, en este caso la intención del voto *John Curtis & Nick Sparrow (1997)*. Para el año de 1997, en las elecciones de Inglaterra, las encuestas volvieron a generar confianza, sin embargo, el método que aplicaron *Gallup e ICM* que fueron las agencias que más se acercaron al resultado, fue mediante vía telefónica *John Curtis (1997)*, un método que les funcionó y que probó tener éxito.

Ahora bien, como la mayoría de los diseños de muestreo, el muestreo por cuotas, parte de la premisa que la muestra debe estar dispersa sobre la población y debe contener la misma fracción de individuos que poseen ciertas características propias de la población. La información censal y de otras encuestas utilizadas para dividir a la población en cierto número de

subpoblaciones mutuamente exclusivas o estratos, además para la determinación o estimación de la fracción de la población en cada uno de los estratos *Stephan et al (1958), Cochran (1984)*.

La muestra es asignada a estratos tales como regiones geográficas, grupos por edad y sexo, nivel socioeconómico de la persona, etc. Esta asignación, va en proporción al tamaño real o estimado de cada estrato o simplemente se utilizan los factores de expansión para balancear la muestra, en caso de que existan diferencias considerables en las proporciones.

Es importante el control de filtros sobre la muestra, en ocasiones son externos a la encuesta, incluso en otros tipos de diseños muestrales. Se debe aclarar que decir que el muestreo por cuota es un muestreo estratificado, es muy arriesgado, porque no es suficientemente poderoso para distinguirlo de otros procedimientos de muestreo *Cochran (1984)*.

El paso final en el muestreo por cuota es la distribución de la muestra entre los enumeradores o encuestadores, asignarles a cada uno el número de personas que deben encuestar de cada estrato. Así sabrán a cuantos hombres y mujeres o cuantas personas con la característica a investigar deben entrevistar. Los enumeradores son libres de elegir a los respondientes con la única condición de que deben llenar o completar su cuota, según los controles de la muestra. Si se contara con la garantía de que los encuestadores seleccionaron a sus respondientes de una forma absolutamente aleatoria dentro de cada estrato o el control de cuota fue aplicado debidamente, el muestreo por cuota podría llamarse muestreo aleatorio estratificado, *Cochran (1984)*.

Hay muchos sentidos que influyen al enumerador para seleccionar a sus encuestados, como preferencias conscientes o inconscientes, un ejemplo podría ser la apariencia del encuestador. La carencia de un control absoluto en esta técnica ha sido siempre reconocida como la debilidad fundamental del muestreo por cuota, tanto por los promotores, como por los opositores al método. Este es el punto coyuntural del muestreo por cuotas, ya que el espíritu de seleccionar las muestras de manera aleatoria desaparece, la mayor parte de la selección queda a disposición de los encuestadores si no es que toda; se menciona la mayor parte, pues deben recaer en supervisores y en las personas que instruyen a los encuestadores *Cochran (1984)*.

En México, las cuotas han sido aplicadas y han funcionado en algunos casos, por ejemplo el Centro de Estudios de Opinión las aplica sólo en la última etapa y en la vía pública, es decir, en la selección de los individuos que habiten en la unidad primaria ó secundaria de muestreo, en las unidades primarias de selección se utiliza un muestreo probabilístico como se ha mencionado salvo en la última etapa de selección, donde los individuos para elegirlos es mediante cuotas; utilizar este método tiene ventajas en los costos y en México facilita la respuesta del entrevistado sobre temas de política o intención de voto. Este método, además de reducir los costos, en la mayoría de los casos, permite llegar a lugares tan alejados con un menor tiempo que un muestreo probabilístico, para el muestreo probabilístico se requiere mayor tiempo y esfuerzo, pues es indispensable localizar a la vivienda o individuo seleccionado en una zona rural donde los traslados de una vivienda a otra son en la mayoría de los casos considerables.

Las personas que habitan en poblaciones rurales, difícilmente opinan sobre la intención del voto en su vivienda, pero en el caso de las ciudades no cambia mucho la situación en nuestro país, de allí la política de implementar cuotas y la entrevista en la vía pública que es una metodología acorde a la realidad nacional al menos en estos momentos. Es importante destacar, que la pregunta filtro al entrevistar una persona, es que ésta habita en la unidad de muestreo seleccionada, que puede ser una Ageb, sección electoral o conglomerado específico, de esta manera se tiene una alta eficacia para disminuir los no respondientes.

Este método tiene limitantes, la principal es que no es probabilístico, pero se le da el tratamiento como si lo fuera para fines prácticos, en la práctica es difícil llevar a cabo al pie de la letra la teoría y lo importante es que este método ha funcionado. Para el caso concreto de Nuevo León, la experiencia, prácticamente es la misma que en el ámbito nacional, en la encuesta preelectoral se utilizaron cuotas por edad y sexo. Las cuotas y las encuestas realizadas en la vía pública funcionaron en la elección previa al 6 de julio de 1997, la experiencia según el coordinador del operativo en campo hacia este método facilita y optimiza el tiempo del levantamiento, pero es importante, que los encuestadores estén altamente capacitados para evitar la preferencia o tendencia de ellos hacia los entrevistados.

Es probable, que en un futuro no lejano, llegue el momento de cambiar el método de cuotas a muestreo probabilístico, pienso, que cuando la población de México se familiarice con las encuestas y pierda el temor de opinar, sin importar si es su casa, su teléfono o cualquier medio de comunicación, ya que en la actualidad, el individuo percibe que no se respetará su anonimato, dejando en entredicho la veracidad de su opinión.

3.- DISEÑOS MUESTRALES DE LAS ENCUESTAS PREELECTORAL, DE SALIDA O POSTVOTO, POST-ELECTORAL Y EL CONTEO RAPIDO EN EL ESTADO DE NUEVO LEON

3.1 INTRODUCCION

En este capítulo, se presentan los diseños muestrales de las encuestas realizadas en el estado de Nuevo León, las descripciones de lo que es una encuesta preelectoral, encuesta de salida, encuesta post-electoral y el conteo rápido se encuentran en la introducción de cada uno de los apartados 3.2, 3.3, 3.4 y 3.5. En el Capítulo dos, se presentaron algunos tipos de muestreo que son la base para el diseño de la muestra de estas encuestas, sin embargo, existen adecuaciones que se deben hacer y ajustar. La notación para los estimadores para proporciones, son los que se utilizan con las variables definidas en el apartado 2.1.5, de esta manera, se presenta la metodología aplicada en cada encuesta electoral. Es en este capítulo donde se hace una aplicación de la teoría a la práctica.

3.2 ENCUESTA PREELECTORAL

3.2.1. INTRODUCCION

Una encuesta preelectoral, es un instrumento que se realiza previo a una elección oficial, mide los parámetros de intención de voto, animadversión (repulsión hacia un partido) y la simpatía por los candidatos principalmente.

Los objetivos para los que se realizan las encuestas preelectorales, varían según el tiempo de anticipación a la elección, del tipo de elección (presidente, gobernador, alcalde, diputados locales o federales, senadores, etc.) y acorde a las necesidades requeridas por el patrocinador. Un ejemplo claro, es ver la posibilidad de triunfo de los partidos políticos sin ser nombrado el candidato de éstos, antes de ser nombrados los candidatos; una encuesta preelectoral serviría para saber quienes son las personas que gozan de mayor popularidad política entre la población en general, para ser posibles pre-candidatos de un partido político determinado y de esta manera conocer las posibilidades que tiene un partido político de ganar una elección. Una

encuesta preelectoral, se convierte en un fuerte instrumento para la toma de decisiones y estrategias para dirigir una campaña con mayores posibilidades de triunfo.

Otro objetivo de la encuesta preelectoral, es medir las necesidades básicas de la población, de esta manera los candidatos o precandidatos a la elección obtienen información para la realización de sus programas, tanto de campaña como de gobierno en caso de ser elegido, cabe mencionar, que esta información también la utilizan los medios de comunicación, la iniciativa privada tanto nacional como extranjera, etc.

Por cuestiones de operatividad, costos y tiempo, se consideró que la encuesta preelectoral a levantarse los días 23 y 24 de junio de 1997 coincidiera con las mismas secciones electorales seleccionadas para la encuesta de salida y el conteo rápido; por lo tanto, el diseño muestral de la encuesta de salida y preelectoral es el mismo, pero el estimador utilizado en este apartado es insesgado a diferencia del de la encuesta de salida, que es de razón. El motivo de que el estimador de la encuesta preelectoral es insesgado a diferencia del de razón y sesgado de la encuesta de salida, es que el denominador corresponde al número de votantes y éste se desconoce, motivo para utilizar el estimador de razón..

Otra diferencia importante, es que la encuesta preelectoral en la etapa final los entrevistados se seleccionaron por cuotas, las cuotas que se manejaron en cada una de las secciones, fueron las que tiene el municipio al que pertenece la sección electoral seleccionada, las cuotas fueron por rangos de:

- a) Edad: 18-24, 25-34, 35-44 y 45 o más años.
- b) Género: Hombres y Mujeres.

Para el operativo de la encuesta preelectoral, se requirió de personal capacitado en gabinete y campo, a todo personal se le capacita mediante cursos breves que contengan los elementos necesarios para poder realizar la función que se le encomienda.

En gabinete:

- un director
- un muestrista
- un analista
- un computólogo
- cuatro procesadores de datos

En campo:

- 1 coordinador
- 2 subcoordinadores
- 3 jefes de grupo
- 45 encuestadores

3.2.2 DISEÑO MUESTRAL

El diseño muestral elaborado para seleccionar la muestra aleatoria de las secciones electorales y seleccionar a los individuos con asignación de cuotas, es un muestreo aleatorio estratificado trietápico en el estrato 7 y un muestreo aleatorio estratificado bietápico de los estratos 1 al 6; en ambos casos con asignación de cuotas sólo en la última etapa. Este diseño tiene el principal propósito de estimar la proporción de votos a favor de cada uno de los partidos políticos en el ámbito estatal en la elección a gobernador del estado de Nuevo León.

El diseño muestral para la encuesta preelectoral, obedece a los aspectos de la encuesta de salida y el conteo rápido que principalmente son tres:

- El primero, la variación en el grado de urbanidad y el tamaño medio de la sección por municipio.

- El segundo, la distribución de la población en el estado.
- El tercero, los ciudadanos con derecho a voto se encuentran aglomerados en secciones electorales y los resultados de la votación se reportan de manera global para cada sección (a nivel casilla); además, la logística y operatividad son aspectos también importantes, dada la naturaleza de la información a registrar, a diferencia de la encuesta de salida y el conteo rápido no existe urgencia en levantar rápidamente las encuestas.

El total de los electores se estratificó en 7 estratos, como se presenta en forma resumida en el cuadro 3-1, la estratificación se hizo como sigue: El estrato 1 contiene al municipio de Monterrey, representante de la capital estatal, tiene 725 secciones con un tamaño medio de 917 electores y concentra al 33.4 % del total de electores reportados en la lista nominal. El estrato 2 corresponde al municipio de Guadalupe, con 279 secciones de tamaño medio de 1250 y representa al 17.5% del total. El estrato 3 representa al municipio de San Nicolás de los Garza, contiene a 238 secciones de tamaño medio 1,174, y un peso relativo del 14%. El estrato 4 está conformado por los municipios de Apodaca y General Escobedo, contiene 152 secciones de tamaño medio 1,113, y un peso relativo de un 8.5%. El estrato 5 corresponde al municipio de San Pedro Garza García, contiene 62 secciones de tamaño medio 1,173, y representa el 3.7% del total de la lista nominal. El estrato 6 contempla el municipio de Santa Catarina, contiene 87 secciones de tamaño medio 1,239, y representa al 5.4% en la lista nominal. Finalmente, el estrato 7 contiene los 44 municipios restantes con 580 secciones de tamaño medio 599 que representan al 17.7% en lista nominal.

Cuadro 3-1 Información de la muestra aplicada por el Centro de Estudios de Opinión de la elección para gobernador del estado de Nuevo León celebrada el 6 de julio de 1997.

Estrato	% de Urb.	# de mun Pob. (muestra)	No. de seccs en la pob.	Tamaño medio de seccs.	% de elect En list. Nom	No. de sec de muestra
1 Monterrey	99%	1 (1)	725	917	33.40%	61
2 Guadalupe	99%	1 (1)	279	1,250	17.50%	23
3 San Nicolás	100%	1 (1)	238	1,174	14.00%	20
4 Apod y Esc.	96%	2 (2)	152	1,113	8.50%	13
5 San Pedro G	99%	1 (1)	62	1,173	3.70%	5
6 Sta Catarina	99%	1 (1)	87	1,239	5.40%	7
7 Otros	65%	44 (10)	580	599	17.50%	49
Total	93%	51 (17)	2,123	937	100%	178

3.2.3 TAMAÑO Y DISTRIBUCION DE LA MUESTRA

El tamaño de muestra fue de 1,800 encuestas distribuidas en las 178 secciones (*ver reporte metodológico preelectoral, 27 de junio de 1997, Centro de Estudios de Opinión*), la cantidad de encuestas por sección electoral fue de un promedio de 10, pero la asignación a cada una de ellas se hizo bajo asignación proporcional con respecto a las que se le asignaron en el estrato correspondiente que también fue proporcional, la cantidad de secciones electorales seleccionadas, como ya se explicó en el apartado 3.2.1, se hizo pensando en la encuesta de salida y el conteo rápido, para poder identificar la sección electoral antes del día de la elección y facilitar el operativo de campo, por consiguiente, según el reporte metodológico de la encuesta de salida y el conteo rápido (*junio de 1997, Centro de Estudios de Opinión*) el tamaño de muestra se determinó en base al siguiente factor:

- En conteos rápidos a nivel estatal realizados por el *Centro de Estudios de Opinión* en Jalisco, Guanajuato, Yucatán, Michoacán y Baja California, donde los tamaños de muestra variaron entre 106 y 155 secciones (con excepción de Jalisco) el error de estimación fue menor a 4 puntos porcentuales en cada uno de los dos o tres partidos mayoritarios.

La distribución de la muestra entre los estratos se hizo de forma proporcional al número de votantes potenciales que pertenecen a cada estrato.

3.2.4 ESTIMADORES PUNTUALES Y POR INTERVALO

En los estratos 1 al 6 existen dos etapas de selección y por tanto el diseño corresponde para esta subpoblación a un muestreo bietápico estratificado. En el caso del último estrato, el 7, existen tres etapas de selección, en la primera se seleccionaron a diez municipios de entre los 44 existentes, dentro de estos municipios se seleccionaron las secciones electorales, las unidades de tercera etapa fueron los individuos seleccionados para la encuesta con asignación de cuotas, para esta subpoblación el diseño corresponde al muestreo trietápico. Se presentan los estimadores puntuales para cada estrato y para el global y los estimadores por intervalo para el global estatal. Para cada sección, antes de salir a campo, se tuvo la certeza de la cantidad de encuestas que se realizarían al momento del operativo, por consiguiente, como se explicó en el apartado 3.2.1 se decidió utilizar el

estimador insesgado trietápico (2.6.15) ó bietápico (2.4.17) según sea el estrato. De esta manera, se presenta el estimador de muestreo estratificado (2.2.6), con la diferencia de que la variable a observarse es dicotómica:

$$\hat{P} = \sum_{t=1}^7 W_t \hat{P}_t, \quad (3.2.1)$$

Notación:

P = proporción verdadera de individuos que poseen cierto atributo de interés, o proporción de electores que votan a favor de un partido específico en el ámbito estatal.

P_t = proporción verdadera de electores que votan a favor de un partido específico en el estrato t ; $t = 1, \dots, 7$.

P_{ni} = proporción verdadera de electores que votan a favor de un partido específico en el estrato t ; $t = 1, \dots, 6$ y la sección i , $i = 1, \dots, n_t$.

\hat{P} = estimador de P

\hat{P}_t = estimador de P_t

\hat{P}_{ti} = estimador de P_{ti}

W_t = peso relativo del estrato t

N_t = número de unidades en la primera etapa del estrato t .

n_t = número de unidades de primera etapa en muestra del estrato t .

M_{ni} = número de unidades en la segunda etapa de la i -ésima unidad primaria del estrato t .

m_{ti} = número de unidades de segunda etapa en muestra de la i -ésima unidad primaria del estrato t .

B_{tij} = número de unidades en la tercera etapa de la j -ésima unidad de segunda etapa de la i -ésima unidad primaria del estrato t .

b_{tij} = número de unidades en muestra de tercera etapa de la j -ésima unidad de segunda etapa de la i -ésima unidad primaria del estrato t .

w_{-t} = esta variable puede consultarse en (2.6.5), es decir:

$$w_{7i} = \frac{Q_{7i}}{Q_7}$$

donde:

$Q_{7i} = M_{7i} \bar{B}_{7i}$ Es decir, la cantidad de electores en el municipio i del estrato 7

$\bar{Q}_{7n} = \frac{\sum_{i=1}^{n_7} Q_{7i}}{N_7}$ Tamaño medio de electores por municipios en el estrato 7.

Para los estratos $t = 1, \dots, 6$ se utilizó el estimador insesgado (2.4.17) de dos etapas:

$$\hat{P}_t = \frac{1}{n_t} \sum_{i=1}^{n_t} \frac{M_{ti}}{\bar{M}_{t n_t}} \hat{P}_{ti}, \quad (3.2.2)$$

el estimador de la varianza quedaría (ver (2.4.19)):
 s_{ip}^2 es el estimador de (2.4.4)

$$\begin{aligned} \hat{V}(\hat{P}_t) &= \left(\frac{1}{n_t} - \frac{1}{N_t} \right) \frac{1}{(n_t - 1)} \sum_{i=1}^{n_t} \left(\frac{M_{ti}}{\bar{M}_{t n_t}} \hat{P}_{ti} - \hat{P}_t \right)^2 \\ &+ \frac{1}{n_t N_t} \sum_{i=1}^{n_t} \left(\frac{1}{m_{ti}} - \frac{1}{M_{ti}} \right) \left(\frac{M_{ti}}{\bar{M}_{t n_t}} \right)^2 s_{ip}^2, \end{aligned} \quad (3.2.3)$$

donde s_{ip}^2 por (2.1.17) queda:

$$s_{ip}^2 = \frac{m_{ti}}{m_{ti} - 1} \hat{P}_{ti} (1 - \hat{P}_{ti}). \quad (3.2.4)$$

En el estrato 7 se utilizó el estimador insesgado (2.6.15) de tres etapas, que se puede visualizar también en (2.6.27):

$$\hat{P}_7 = \frac{1}{n_7} \sum_{i=1}^{n_7} w_{7i} \frac{1}{m_{7i}} \sum_{j=1}^{m_{7i}} \frac{B_{7ij}}{B_{7i}} \hat{P}_{7ij} \quad (3.2.5)$$

Para entender mejor este estimador se explica:

$$\hat{p}_{7ij} = \frac{\text{entrevistados a favor de un partido determinado en la seccion } i \text{ del municipio } j \text{ del estrato } 7}{\text{entrevistados totales en la seccion } i \text{ del municipio } j \text{ del estrato } 7} \quad (3.2.6)$$

Es decir, se genera un estimador de la sección j del municipio i del estrato 7, de esta manera se procede a crear el estimador del municipio i del estrato 7, que es el estimador (3.2.7):

$$\hat{p}_{i,} = \frac{1}{m_{7i}} \sum_{j=1}^{m_{7i}} \frac{B_{7ij}}{B_{7i}} \hat{p}_{7ij}, \quad (3.2.7)$$

al obtener el estimador para el municipio i del estrato 7, entonces queda:

$$\hat{p}_7 = \frac{1}{n_7} \sum_{i=1}^{n_7} w_{7i} \hat{p}_{i,}, \quad (3.2.8)$$

esto básicamente está repitiendo el proceso anterior, después de generar un estimador por municipio, se procede a darle el peso que estos tienen respecto a los otros municipios seleccionados en el estrato 7.

El estimador de la varianza del estimador (3.2.5) es (2.6.17):

$$\mathcal{V}(\hat{p}_7) = \left(\frac{1}{n_7} - \frac{1}{N_7} \right) s_{7bp}^2 + \frac{1}{n_7 N_7} \sum_{i=1}^{N_7} w_{7i}^2 \left(\frac{1}{m_{7i}} - \frac{1}{M_{7i}} \right) s_{7ip}^2 + \frac{1}{n_7 N_7} \sum_{i=1}^{N_7} \frac{w_{7i}^2}{m_{7i} M_{7i}} \sum_{j=1}^{M_{7i}} \frac{B_{7ij}^2}{B_{7i}^2} \left(\frac{1}{b_{7ij}} - \frac{1}{B_{7ij}} \right) s_{7ijp}^2, \quad (3.2.9)$$

Así la varianza del estimador (3.2.1) es (2.2.8):

$$\mathcal{V}(\hat{p}) = \sum_{i=1}^7 W_i^2 \mathcal{V}(\hat{p}_i), \quad (3.2.10)$$

y el estimador de la varianza del estimador (3.2.10) es (3.2.11):

$$\hat{\mathcal{V}}(\hat{p}) = \sum_{i=1}^7 W_i^2 \hat{\mathcal{V}}(\hat{p}_i), \quad (3.2.11)$$

así el intervalo de confianza con una probabilidad de 95% para aquellas proporciones entre 0.2 y 0.8, se aproximan con la expresión siguiente:

$$p \pm 1.96 \sqrt{\hat{p}(1-\hat{p})} \quad (3.2.12)$$

3.3 ENCUESTA DE SALIDA O POSTVOTO

3.3.1. INTRODUCCION

Una encuesta de salida, es aquella que se realiza a los electores después de haber emitido su voto en la casilla electoral respectiva, de allí el sinónimo postvoto.

La metodología que se sigue en el operativo en campo, es nombrar un responsable por sección electoral seleccionada y a éste se le asigna una cuota efectiva por cumplir, es decir, cierta cantidad de electores que emitieron su voto y respondieron favorablemente a la pregunta que se le hizo. De existir más de una casilla se procedió a encuestar en la básica.

Para el operativo de la encuesta de salida y el conteo rápido, se requirió de personal capacitado y experimentado en gabinete y campo, a todo personal se les capacita mediante cursos breves que contengan los elementos necesarios para poder realizar la función que le requiere, además, cabe destacar que la experiencia es adquirida en procesos de encuestas electorales para que el personal sepa afrontar los inconvenientes que se requieran:

En gabinete:

- un director
- un muestrista
- un analista
- un programador
- un coordinador
- cuatro capturistas
- ocho telefonistas

En campo:

- un coordinador
- cuatro subcoordinadores
- diecisiete jefes de grupo
- ciento setenta y ocho encuestadores

La jornada electoral, comenzó con la apertura de las casillas a las 8:00 y cerraron a las 18:00hrs del 6 de julio de 1997, en postvoto se realizaron 3 cortes y su respectivo acumulado, los cortes se realizaron a las 10:30, 14:30 y 18:00hrs, de esta manera se puede conocer la intención del voto hacia los partidos en los diferentes horarios y la evolución del mismo, el criterio para seleccionar los individuos de última etapa es seleccionando los electores después de emitir su voto con cierta distribución de cuota y que éstos sean efectivos (10 encuestas), se pide que el encuestador observe la distribución de los votantes que acuden a las urnas y las distribuya de tal manera que sea un hombre por cada mujer, salvo que se observe una marcada diferencia en la participación por género y no se consideró edad.

Un corte parcial es aquel que sólo da información en el tiempo en que la encuesta fue levantada, mas no del proceso en su conjunto, de esta manera, se puede conocer por cual partido se votó en la mañana, por cual al mediodía y por último en la tarde.

Un corte acumulado, es aquel que dice la intención del voto desde el inicio de la jornada hasta el momento del corte de postvoto, aquí se está en posibilidad de ver la tendencia a través del tiempo. Se puede realizar una curva, que permita visualizar el comportamiento a través del tiempo de la intención del voto.

Para postvoto, se planeó que en la encuesta preelectoral se detectaran las secciones electorales en las que se iban a aplicar la encuesta de salida, de esta manera, se obtuvo una ventaja en la cuestión logística para una mayor rapidez del operativo en campo y se detectaron con anterioridad los contratiempos. La metodología, similar que la encuesta preelectoral asimiló experiencia, se seleccionaron los mismos puntos de muestreo bajo el mismo método, la diferencia fue que en la encuesta preelectoral los encuestadores sabían cuantos casos iban a entrevistar por sección, en la encuesta de salida llevaron la consigna de levantar 10 casos efectivos por corte (se requirió de un promedio de 12.25 encuestas por corte para completar la cuota). La fecha de levantamiento de esta encuesta fue el día 6 de julio de 1997 (día de la elección para gobernador).

3.3.2 DISEÑO MUESTRAL

El diseño muestral elaborado para seleccionar la muestra aleatoria de las secciones electorales y seleccionar a los individuos, es un muestreo aleatorio estratificado trietápico en el estrato 7 y del estrato 1 al estrato 6 un muestreo aleatorio estratificado bietápico. Este diseño es el mismo que se expone en el apartado 3.2.3 y tiene como principal propósito, estimar la proporción de votos a favor de los partidos políticos mayoritarios de la elección para gobernador.

En esta encuesta fue necesario levantar rápidamente las encuestas, ya que, la logística y operatividad son aspectos importantes dada la naturaleza de la información y la rapidez para concentrarla, para que pueda tenerse información oportuna. Para la logística de campo se necesitó personal experimentado y capacitado tanto en gabinete, que es el centro de acopio donde se recibe la información global para su proceso y en campo, que básicamente son los encuestadores.

3.3.3 TAMAÑO Y DISTRIBUCION DE LA MUESTRA

La distribución de la muestra se presenta en el cuadro 3-1, donde se puede verificar que el número de estratos son 7 al igual que la encuesta preelectoral.

El tamaño de muestra se explicó en el apartado 3.2.3, se hace referencia sobre la base de la experiencia, de esta manera se espera un total de $30 \times 178 = 5,340$ encuestas efectivas a lo largo de la jornada electoral, que requirieron de un total de 6,542 para completar la cuota. Se completaron un total de 4,826 encuestas efectivas de un total de 6,227.

3.3.4 ESTIMADORES PUNTUALES Y POR INTERVALO

Los estimadores que aquí se presentan son adaptados a la notación de esta tesis, esto para mayor comprensión del lector, si el lector desea consultar la notación publicada, entonces véase reporte metodológico encuesta de salida y conteo rápido (*junio de 1997, Centro de Estudios de Opinión*).

La notación a utilizar es la misma que la de la encuesta preelectoral en el apartado 3.2.4.

El estimador global a utilizar de un muestreo aleatorio estratificado equivalente al estimador (2.2.6) es el siguiente:

$$\hat{p} = \sum_{i=1}^7 W_i \hat{p}_i, \quad (3.3.1)$$

Para los estratos $t = 1, \dots, 6$, tenemos un muestreo bietápico donde el estimador aplicado es el de razón equivalente al (2.4.20) de esta tesis: donde:

$$\hat{p}_t = \frac{\sum_{i=1}^{n_t} M_{it} \hat{p}_{it}}{\sum_{i=1}^{n_t} M_{it}}, \quad (3.3.2)$$

$$\hat{p}_{it} = \frac{\text{entrevistados a favor de un partido en la seccion } i \text{ del estrato } t}{\text{entrevistados totales en la seccion } i \text{ del estrato } t}. \quad (3.3.3)$$

Una aproximación del estimador de la varianza es (2.4.22):

$$\begin{aligned} \mathcal{V}(\hat{p}_t) = & \left(\frac{1}{n_t} - \frac{1}{N_t} \right) \frac{1}{(n_t - 1)} \sum_{i=1}^{n_t} \left(\frac{M_{it}}{M_{t..}} \right)^2 (\hat{p}_{it} - \hat{p}_t)^2 \\ & + \frac{1}{n_t N_t} \sum_{i=1}^{n_t} \left(\frac{1}{m_{it}} - \frac{1}{M_{it..}} \right) \left(\frac{M_{it}}{M_{it..}} \right)^2 s_{iip}^2, \end{aligned} \quad (3.3.4)$$

s_{iip}^2 es el estimador de (2.4.4) pero por (2.1.17) queda:

$$s_{iip}^2 = \frac{m_{it}}{m_{it} - 1} \hat{p}_{it} (1 - \hat{p}_{it}) \quad (3.3.5)$$

Para el estrato 7 se utilizó un modelo de muestreo trietápico, la manera de construirlo se presenta a continuación.

$$\hat{P}_{7ij} = \frac{\text{entrevistados a favor de un partido determinado en la seccion i del municipio j del estrato k}}{\text{entrevistados totales en la seccion i del municipio j del estrato k}} \quad (3.3.6)$$

$$\hat{P}_{7i} = \frac{\sum_{j=1}^{b_{7y}} \frac{B_{7ij}}{b_{7y}} \hat{P}_{7ij}}{\sum_{j=1}^{b_{7y}} B_{7ij}} \quad (3.3.7)$$

$$\hat{P}_7 = \frac{\sum_{i=1}^{m_{7i}} \frac{M_{7i}}{m_{7i}} \hat{P}_{7i}}{\sum_{i=1}^{m_{7i}} M_{7i}} \quad (3.3.8)$$

El estimador de la varianza de un muestreo trietápico de razón es el estimador (2.6.38).

$$\begin{aligned} \varphi(\hat{P}_7) &= \left(\frac{1}{n_7} - \frac{1}{N_7} \right) \left(s_{7bp}^2 - 2\hat{P}_7 s_{7bpx} + \hat{P}_7^2 s_{7bx}^2 \right) \\ &+ \frac{1}{n_7 N_7} \sum_{i=1}^{n_7} w_{7i}^2 \left(\frac{1}{m_{7i}} - \frac{1}{M_{7i}} \right) \left(s_{7ip}^2 - 2\hat{P}_7 s_{7ipx} + \hat{P}_7^2 s_{7ix}^2 \right) \\ &+ \frac{1}{n_7 N_7} \sum_{i=1}^{n_7} \frac{w_{7i}^2}{m_{7i} M_{7i}} \sum_{j=1}^{m_{7i}} \frac{B_{7ij}^2}{B_{7i}^2} \left(\frac{1}{b_{7ij}} - \frac{1}{B_{7ij}} \right) \left(s_{7ijp}^2 - 2\hat{P}_7 s_{7ijpx} + \hat{P}_7^2 s_{7ijx}^2 \right) \end{aligned} \quad (3.3.9)$$

Así la varianza del estimador (3.3.1) es (2.2.8):

$$V(\hat{P}) = \sum_{i=1}^7 W_i^2 V(\hat{P}_i) \quad (3.3.10)$$

el estimador de la varianza del estimador es:

$$\varphi(\hat{P}) = \sum_{i=1}^7 W_i^2 \varphi(\hat{P}_i) \quad (3.3.11)$$

y el intervalo de confianza para aquellas proporciones entre 0.2 y 0.8, se aproximan con la expresión siguiente:

$$P \pm 1.96 \sqrt{\hat{p}(\hat{p})} \quad (3.3.12)$$

3.4 CONTEO RAPIDO

3.4.1. INTRODUCCION

Un conteo rápido, es una muestra que consiste en recabar la cantidad de votos hacia cada uno de los partidos políticos el día de la elección después que la casilla electoral haya cerrado y los funcionarios de éstas hayan publicado oficialmente sus resultados en las afueras de ésta. El conteo rápido, no es una encuesta, ya que en ningún momento se entrevista a los electores.

Entre los objetivos principales de un conteo rápido, es el respaldo de un resultado oficial, de esta manera, una elección es más limpia; otro objetivo claro y principal es el conocer el resultado para poder determinar el mismo día de la elección quien obtuvo el triunfo electoral de la jornada. En el caso concreto de Nuevo León para saber con anticipación a los resultados oficiales quien será el próximo gobernador.

3.4.2 DISEÑO MUESTRAL

El diseño muestral elaborado para seleccionar la muestra aleatoria de secciones para el conteo rápido, corresponde a un diseño de submuestreo estratificado en el estrato 7 y un muestreo estratificado por conglomerados del estrato 1 al estrato 6. Este diseño y el tamaño de la muestra tienen como objetivo único estimar la proporción de votos a favor de cada uno de los partidos políticos en el ámbito estatal de la elección para gobernador.

La diferencia con los apartados 3.2.2 y 3.1.2 en el diseño muestral es que el conteo no es encuesta y en los apartados anteriores se realizaron encuestas. En el caso del conteo rápido se registra la información de la votación total de la sección electoral, entendiéndose ésta como el total de votación en las casillas electorales que componen la sección, además como no se realizó etapa de selección para los individuos, el conteo rápido consta de una etapa de selección menos que las encuestas.

3.4.3 TAMAÑO Y DISTRIBUCION DE LA MUESTRA

El tamaño de la muestra fue de 178 secciones electorales, el tamaño de muestra también contempla un porcentaje estimado del 10% de no registro de resultado de secciones electorales seleccionadas en la muestra. El no registro se debe a que algunas de las secciones electorales emiten el resultado oficial después de las 22:30 hrs., o bien a que se presentan otros contratiempos que impiden registrar el resultado oficial con agilidad, el tamaño de muestra fue descrito en el apartado 3.2.3. La operatividad de campo y calidad de la información se garantizó con un encuestador responsable de registrar y reportar los resultados de una sección electoral, supervisados por un coordinador general.

3.4.4 ESTIMADORES PUNTUALES Y POR INTERVALO

Al igual que las encuestas de salida y preelectoral se hicieron 6 estratos donde sólo existe una sola etapa de selección y por tanto el diseño corresponde para esta subpoblación a un muestreo por conglomerados estratificado. En el caso del último estrato, el 7, existen dos etapas de selección, en la primera se seleccionaron a diez municipios de entre los 44 existentes, dentro de éstos se seleccionaron las secciones representantes del estrato, por consiguiente, corresponde a un muestreo bietápico. Se presentan los estimadores puntuales para cada estrato y para el global, los estimadores por intervalo para el global estatal y la notación que se usará en este apartado es la misma que se utiliza en el capítulo.

Se presenta el estimador global a utilizar de un muestreo aleatorio estratificado como primer paso, mostrado en (2.2.6):

$$\hat{P} = \sum_{t=1}^7 W_t \hat{P}_t. \quad (3.4.1)$$

Para los estratos $t = 1, \dots, 6$, se tiene un muestreo por conglomerados de una etapa, donde el estimador aplicado es el de razón mostrado en (2.3.7) de esta tesis. De esta manera, el estimador de los estratos 1 a 6 es:

$$\begin{aligned} \hat{P}_t &= \sum_{i=1}^{n_t} \frac{M_{it}}{\sum_{i=1}^{n_t} M_{it}} P_{it} = \frac{\sum_{i=1}^{n_t} M_{it} \sum_{j=1}^{M_{it}} y_{ij}}{\sum_{i=1}^{n_t} M_{it}} = \frac{\sum_{i=1}^{n_t} \sum_{j=1}^{M_{it}} y_{ij}}{\sum_{i=1}^{n_t} M_{it}} \\ &= \frac{\text{votos a favor de un partido en el estrato } t}{\text{votos totales en el estrato } t}, \end{aligned} \quad (3.4.2)$$

al menos para los estratos 1 a 6 se ha probado dentro de la construcción del estimador (3.4.2) que es autoponderado para el estimador de cada estrato, lo que nos facilita los procesos de datos en cómputo, esto fue posible gracias a que se censó en el conglomerado.

Una aproximación de la varianza del estimador (3.4.2) es (2.3.8):

$$V(\hat{p}_i) = \left(\frac{1}{n_i} - \frac{1}{N_i} \right) S_{ir}^2, \quad (3.4.3)$$

donde S_{ir}^2 es (2.3.9), si se sustituye (2.3.9) por (2.3.10) se obtiene una estimación de la varianza:

$$v(\hat{p}_i) = \left(\frac{1}{n_i} - \frac{1}{N_i} \right) s_{ir}^2, \quad (3.4.4)$$

y (2.3.10) es:

$$s_{ir}^2 = \sum_{i=1}^{n_i} \left(\frac{M_{ii}}{\bar{M}_{n_i}} \right)^2 \frac{(\hat{p}_{ii} - \hat{p}_i)^2}{n_i - 1}. \quad (3.4.5)$$

Para el estimador del estrato 7, se utilizó un submuestreo aleatorio correspondiente a un estimador de razón (2.4.20), la manera de construirlo es la siguiente:

$$\hat{p}_{7i} = \frac{\text{entrevistados a favor de un partido en la seccion } i \text{ del estrato } 7}{\text{entrevistados totales en la seccion } i \text{ del estrato } 7} \quad (3.4.6)$$

una aproximación del estimador de la varianza es (2.4.22):

$$v(\hat{p}_7) = \left(\frac{1}{n_7} - \frac{1}{N_7} \right) \frac{1}{(n_7 - 1)} \sum_{i=1}^{n_7} \left(\left(\frac{M_{7i}}{\bar{M}_{7n_7}} \right)^2 (\hat{p}_{7i} - \hat{p}_7)^2 \right) + \frac{1}{n_7 N_7} \sum_{i=1}^{n_7} \left(\frac{1}{m_{7i}} - \frac{1}{M_{7i}} \right) \left(\frac{M_{7i}}{\bar{M}_{7n_7}} \right)^2 s_{7ip}^2, \quad (3.4.7)$$

donde s_{7ip}^2 es el estimador (2.4.4) pero por (2.1.16) queda:

$$s_{7ip}^2 = \frac{m_{7i}}{m_{7i} - 1} \hat{p}_{7i} (1 - \hat{p}_{7i}), \quad (3.4.8)$$

Así la varianza del estimador (3.4.1) es (2.2.8):

$$V(\hat{p}) = \sum_{i=1}^I W_i^2 V(p_i) \quad (3.4.9)$$

el estimador de la varianza del estimador es:

$$\hat{v}(\hat{p}) = \sum_{i=1}^I W_i^2 \hat{v}(p_i) \quad (3.4.10)$$

y el intervalo de confianza para aquellas proporciones entre 0.2 y 0.8, se aproximan con la expresión siguiente:

$$\hat{p} \pm 1.96 \sqrt{\hat{v}(\hat{p})} \quad (3.4.11)$$

3.5 ENCUESTA POST-ELECCIONAL

3.5.1. INTRODUCCION

Una encuesta post-electoral es aquella que se realiza posterior a una elección oficial para conocer los motivos que llevaron a los votantes a decidir su preferencia electoral, generalmente se considera post-electoral si ésta se realiza posterior al día de la elección, el tiempo puede variar y los criterios para considerarla post-electoral son poco claros.

Entre los objetivos principales se encuentra el preguntar la razón de la decisión de la preferencia electoral de cada uno de los partidos políticos, cuando decidió votar por el partido que eligió, conocer los motivos de los que no fueron a votar por que lo hicieron.

Valorar el impacto de las campañas sobre la decisión del voto, las sugerencias al candidato y la valuación de la jornada electoral.

Un punto importante es saber que hubiese sucedido si los candidatos para gobernador hubieran sido diferentes, si los electores simpatizaron o no con él, que tanto influyó el candidato o el partido en el momento de la decisión final.

Saber si las personas pertenecen a alguna organización social, tales como la iglesia, sindical, deportiva, empresarial, etc. y si ésta influyó en la decisión final sobre su preferencia electoral.

Valorar la imagen que el presidente de México tuvo en el estado hacia el apoyo de la democracia, si es aprobatoria o no, además de estar en condiciones de ponerle una calificación de su apoyo a la democracia en el estado de Nuevo León en la elección para gobernador.

Una pregunta interesante es conocer la aceptación de realización de encuestas preelectorales en el estado y sus diferentes razones, así como saber si la población se enteró de los resultados de alguna encuesta y los problemas que la próxima administración ha de enfrentar.

En esta encuesta post-electoral se utiliza un muestreo trietápico-bietápico estratificado, con un estimador insesgado, la diferencia radica de los muestreos anteriores en que los estratos se definieron por distrito electoral federal, se eligieron federales porque son 11 y la cantidad de locales son 26, el objetivo a medir es el ámbito estatal, por lo tanto no nos afecta si los distritos son locales o federales por tratarse solamente de diferentes formas de partir el universo. Existen distritos en los que se aplicó un muestreo bietápico, mientras que en otros se aplicó un trietápico; a pesar que son 11 distritos electorales, el número de estratos que se definieron son 12, en el distrito 3 existen dos municipios y se decidió estratificar. La encuesta post-electoral se hizo para conocer las razones que llevaron a votar por determinado partido, evaluar la percepción de la gente respecto a la jornada electoral, la fecha de levantamiento de la encuesta fueron los días 13 y 14 de julio de 1997.

Para el operativo de la encuesta post-electoral, se requirió de personal capacitado en gabinete y campo, a todo personal se le capacita mediante cursos breves que contengan los elementos necesarios para poder realizar la función que le requiere.

En gabinete:

- un director
- un muestrista
- un analista
- un computólogo
- cuatro procesadores de datos

En campo:

- un coordinador
- un subcoordinador
- dos jefes de grupo
- treinta encuestadores

3.5.2 DISEÑO MUESTRAL

El diseño muestral de este estudio se hizo bajo el principal parámetro de por quien votó, este reactivo desde luego que se preguntó en la encuesta, sin embargo, otro objetivo importante es medir porque los electores votaron por ese partido. En esta encuesta, no existe ningún tipo de presión con respecto al tiempo, la cantidad de reactivos que se hicieron en esta encuesta fueron 27 a diferencia de la preelectoral que fueron 25 reactivos.

Los estratos se definieron por los 11 distritos electorales federales y el distrito 3 en dos estratos quedando así un total de 12, dentro de cada distrito electoral se seleccionaron las secciones electorales de manera simple aleatoria; seleccionada ésta, se procedió a seleccionar los individuos que habitan dentro de ésta con la asignación de cuotas por género y edad con la libre decisión del encuestador siempre y cuando el encuestador habite en la sección electoral seleccionada. En este diseño se utiliza el estimador insesgado bietápico (2.4.17) para los distritos: 3 con su partición de los municipios General Escobedo y San Nicolás de los Garza; distritos 4; 5; 6; 7; 8; 10; 11; de esta manera, en 9 de los 12 estratos se aplicó un modelo estratificado bietápico y para los distritos 1, 2 y 9 un modelo estratificado trietápico.

3.5.3 TAMAÑO Y DISTRIBUCION DE LA MUESTRA

El tamaño de muestra fue de 1,209 casos *reporte metodológico, Encuesta Post-electoral (julio de 1997, Centro de Estudios de Opinión)*, bajo un muestreo aleatorio simple con una confianza de 95% y un error estándar de $\pm 2.8\%$, si $P = 0.5$. Se visitaron un total de 80 de las 2,123 secciones electorales que representan el 3.77% del total de las secciones del estado, además se visitaron un total de 14 de los 51 municipios que existen en el estado de Nuevo León.

La cantidad de encuestas realizadas por estrato se hizo bajo el criterio de asignación proporcional. Definidas la cantidad de encuestas a realizar en cada uno de los estratos, en los estratos 3, 4, 5, 6, 7, 8, 9, 11 y 12 se procedió a seleccionar una cantidad de secciones electorales donde la cantidad de encuestas promedio por sección sea de 15, por ejemplo, en el estrato 3 donde

se realizaron 54 encuestas, al dividir 54/15 nos quedan aproximadamente 3 secciones electorales que son las que se van a seleccionar en el estrato 3.

La manera de seleccionar las secciones electorales se hizo asignándoles números del 1 hasta las que hayan en el estrato, posteriormente se generaron tantos números aleatorios como secciones electorales a seleccionar, los números aleatorios entre 0 y 1, éstos se multiplicaron por la cantidad de secciones electorales que existen en el estrato, de esta manera se seleccionaron las secciones. Seleccionadas las secciones, se procedió a asignarles una cantidad proporcional con respecto al resto de las secciones electorales dentro del estrato o municipio según sea el caso. En el cuadro 3-2 se presenta la información sobre el diseño de la muestra de la encuesta post-electoral.

Cuadro 3-2 Distribución de la muestra en la encuesta post-electoral en el estado de Nuevo León levantada por el Centro de Estudios de Opinión los días 12 y 13 de julio de 1997

ESTRATO	DIST	MUNICIPIOS EN MUESTRA	MUNICIPIOS POR ESTRATO	SECCIONES POR ESTRATO	SECCS EN MUESTRA	ENCUESTAS	% LISTA NOMINAL
1	1	LAMPAZOS DE NARANJO, SAN PEDRO GARZA GARCIA, SANTA CATARINA	7	200	9	128	10.50%
2	2	AGUALEGUAS, APODACA, CADEREYTA JIMENEZ, CHINA	26	316	9	134	11.08%
3	3	SAN NICOLAS DE LOS GARZA	1	87	3	54	4.87%
4	3	GENERAL ESCOBEDO	1	64	3	42	3.74%
5	4	SAN NICOLAS DE LOS GARZA	1	151	8	120	9.16%
6	5	MONTERREY	1	133	8	105	8.80%
7	6	MONTERREY	1	132	6	97	7.94%
8	7	MONTERREY	1	291	6	105	8.77%
9	8	GUADALUPE	1	135	7	101	8.40%
10	9	GENERAL TERAN, LINARES, SANTIAGO	14	301	7	121	9.72%
11	10	MONTERREY	1	169	7	94	7.90%
12	11	GUADALUPE	1	144	7	108	9.12%
TOTAL				2123	80	1209	100.00%

Para los estratos 1, 2 y 10 primero se seleccionaron los municipios mostrados en el cuadro 3-2, después se procedió a seleccionar las secciones electorales de igual manera que el resto de los estratos.

3.5.4 ESTIMADORES PUNTUALES Y POR INTERVALO

El estimador global corresponde a un muestreo aleatorio estratificado presentado en (2.2.6), la notación es la misma definida para este capítulo:

$$\hat{p} = \sum_{t=1}^{12} W_t \hat{p}_t. \quad (3.5.1)$$

Un estimador de la varianza de (3.5.1) según (2.2.8) es:

$$\hat{v}(\hat{p}) = \sum_{t=1}^{12} W_t^2 \hat{v}(\hat{p}_t), \quad (3.5.2)$$

el estimador de $t = 1, 2$ y 10 se presentan a continuación:

En el estrato t se utilizó el estimador insesgado (2.6.15) de tres etapas, que se puede visualizar en (2.6.27):

$$\hat{p}_t = \frac{1}{n_t} \sum_{i=1}^{n_t} w_{it} \frac{1}{m_{it}} \sum_{j=1}^{m_{it}} \frac{B_{ij}}{\bar{B}_{it}} \hat{p}_{ij}, \quad (3.5.3)$$

Para entender mejor este estimador explicamos:

$$\hat{p}_{ij} = \frac{\text{entrevistados a favor de un partido determinado en la seccion } i \text{ del municipio } j \text{ del estrato } t}{\text{entrevistados totales en la seccion } i \text{ del municipio } j \text{ del estrato } t} \quad (3.5.4)$$

Es decir, se hace un estimador de la sección j del municipio i del estrato t y se procede a crear el estimador del municipio i del estrato t , que es el estimador:

$$\hat{p}_n = \frac{1}{m_n} \sum_{j=1}^{m_n} \frac{B_{ij}}{\bar{B}_n} \hat{p}_{ij}, \quad (3.5.5)$$

al obtener el estimador para el municipio i del estrato t , queda:

$$\hat{p}_t = \frac{1}{n_t} \sum_{i=1}^{n_t} w_{it} \hat{p}_{it}, \quad (3.5.6)$$

esto básicamente está repitiendo el proceso anterior, después de generar un estimador por municipio, se procede a darle el peso que éstos tienen respecto a los otros municipios seleccionados en el estrato t .

Un estimador de la varianza del estimador (3.5.3) nos queda:

$$\hat{v}(\hat{p}_t) = \left(\frac{1}{n_t} - \frac{1}{N_t} \right) s_{itp}^2 + \frac{1}{n_t N_t} \sum_{i=1}^{n_t} w_{it}^2 \left(\frac{1}{m_{it}} - \frac{1}{M_{it}} \right) s_{itp}^2 + \frac{1}{n_t N_t} \sum_{i=1}^{n_t} \frac{w_{it}^2}{m_{it} M_{it}} \sum_{j=1}^{m_{it}} \frac{B_{ij}^2}{\bar{B}_{it}^2} \left(\frac{1}{b_{ij}} - \frac{1}{\bar{B}_{ij}} \right) s_{itp}^2, \quad (3.5.7)$$

Para el resto de los estratos se tiene un muestreo bietápico, $t = 3, 4, 5, 6, 7, 8, 9, 11$ y 12 y se utiliza el estimador insesgado (2.4.17) de dos etapas:

$$\hat{p}_t = \frac{1}{n_t} \sum_{i=1}^{n_t} \frac{M_{it}}{M_{tN_t}} \hat{p}_{it} \quad (3.5.8)$$

Simplemente lo que está diciendo de este estimador es que se genera un estimador para proporciones por sección electoral, posterior a eso se le da el peso específico a cada estimador por sección con respecto a las demás secciones seleccionadas en el estrato y de allí se genera (3.5.8).

Un estimador de la varianza del estimador (3.5.8) es (2.4.19):

$$\hat{v}(\hat{p}_t) = \left(\frac{1}{n_t} - \frac{1}{N_t} \right) \frac{1}{(n_t - 1)} \sum_{i=1}^{n_t} \left(\frac{M_{it}}{M_{tN_t}} \hat{p}_{it} - \hat{p}_t \right)^2 + \frac{1}{n_t N_t} \sum_{i=1}^{n_t} \left(\frac{1}{m_{it}} - \frac{1}{M_{it}} \right) \left(\frac{M_{it}}{M_{tN_t}} \right)^2 s_{itp}^2, \quad (3.5.9)$$

donde s_{itp}^2 es (2.4.4) pero por (2.1.17) queda:

$$s_{itp}^2 = \frac{m_{it}}{m_{it} - 1} \hat{p}_{it} (1 - \hat{p}_{it}) \quad (3.5.10)$$

así el intervalo de confianza con una probabilidad de 95% para aquellas proporciones entre 0.2 y 0.8, se aproximan con la expresión siguiente:

$$p \pm 1.96 \sqrt{p(1-p)} \quad (3.5.11)$$

**ESTA TESIS NO DEBE
SALIR DE LA BIBLIOTECA**

4.- PRESENTACION DE RESULTADOS

4.1. INTRODUCCION

En este capítulo mostramos los resultados de la última encuesta preelectoral de las dos que se realizaron por el Centro de Estudios de Opinión; así se tiene que en el punto 4.2 se presentan los resultados generales de la encuesta; esto es: su intención de voto hacia los diferentes partidos políticos, así como su animadversión, se analizaron sólo algunos cruces de variables, los que consideramos a nuestro juicio los más representativos del tema, y las cuotas de esta encuesta. En el punto 4.3 se presenta la encuesta de salida o postvoto, donde se establece sólo la intención de voto por corte parcial y la de corte acumulado. En el punto 4.4 se presenta el conteo rápido, la curva de evolución del voto en el Estado de Nuevo León conforme fueron llegando los resultados por la cantidad de secciones electorales. En el punto se presentan los resultados oficiales, lo cual nos da un momento de reflexión y análisis sobre el trabajo realizado en comparación con el oficial que el órgano competente presentó. En el punto 4.6 se presenta la encuesta post-electoral, donde se analizan las razones mostradas por el electorado: ¿Por qué votaron por cierto partido político?.

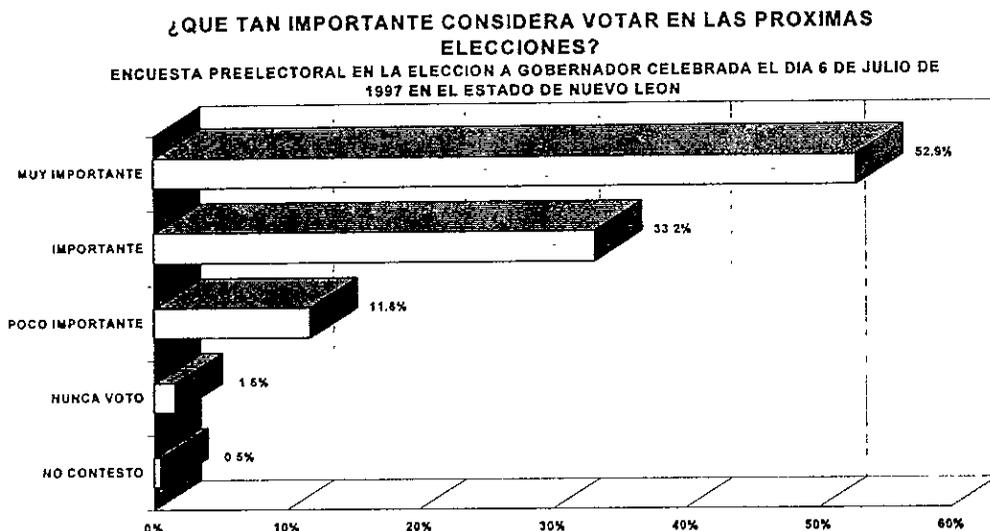
4.2. ENCUESTA PREELECTORAL

4.2.1 RESULTADOS DE LA ENCUESTA PREELECTORAL

A continuación se presentan las gráficas del estudio, así como los comentarios que vertimos al respecto de cada una de ellas, con el fin de tener un referente que nos posibilite el análisis y la comprensión del trabajo que en la presente tesis nos ocupa.

Como punto de inicio se estimó que el 94.1% de la población tiene credencial para votar con fotografía, que le permite votar en el Estado de Nuevo León; dicha población lógicamente es mayor de 18 años de edad. Según este resultado, se puede apreciar la importancia que para la población adulta representa el estar empadronado.

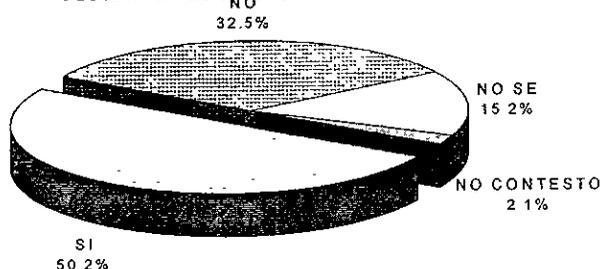
Gráfico 4.2.A Importancia de la población para votar en la próxima elección



La gráfica mide la importancia que para la población representa el acudir a votar en las elecciones del 6 de julio de 1997 para gobernador del Estado de Nuevo León, cabe destacar que la población del Estado mayoritariamente considera ir a votar muy importante, prueba de ello está en que el 52.9% expresa que es muy importante, mientras que el 33.2% menciona que es importante, sólo un 11.8% de la población considera poco importante ir a votar, el 1.6%, como se puede ver en la gráfica, nunca vota y sólo el 0.5% no contestó; de esta manera podemos concluir que la gente en el Estado de Nuevo León tiene la convicción democrática o al menos de la necesidad de acudir a votar; las personas que en el proceso electoral sean elegidas en los puestos que se postularon deben considerar el sentido democrático de la población, de ellas depende en gran medida el próximo rumbo del Estado en el ámbito político, social y económico.

Gráfico 4.2.B Creencia de la población sobre la honestidad de las elecciones

¿CREE QUE LAS ELECCIONES PARA GOBERNADOR DE NUEVO LEON SERAN LIMPIAS Y HONESTAS?
ENCUESTA PREELECTORAL EN LA ELECCION A GOBERNADOR CELEBRADA EL 6 DE JULIO DE 1997



En la gráfica 4.2.B se observa que la mayoría de la ciudadanía cree que las elecciones serán limpias, esto permite a la ciudadanía tener una mayor disposición hacia la intención de acudir a votar. El gráfico 4.2.C nos mide la popularidad de los candidatos para gobernador, donde se puede corroborar que el candidato más conocido es José Natividad del PRI, sin embargo, la diferencia con Fernando Canales del PAN no es considerablemente mayor como para decir con cierto nivel de certidumbre que el candidato del PRI será el próximo gobernador.

Gráfico 4.2.C Nivel de conocimiento de la ciudadanía hacia los candidatos

¿SABE EL NOMBRE DE LOS CANDIDATOS A LA GUBERNATURA DEL ESTADO DE NUEVO LEON?
ENCUESTA PREELECTORAL EN LA ELECCION A CELEBRARSE EL 6 DE JULIO DE 1997

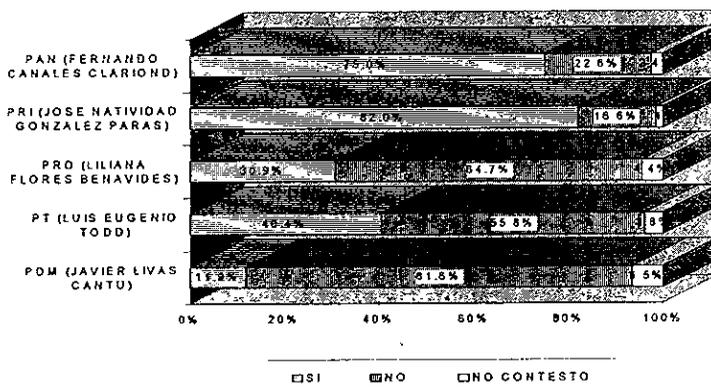
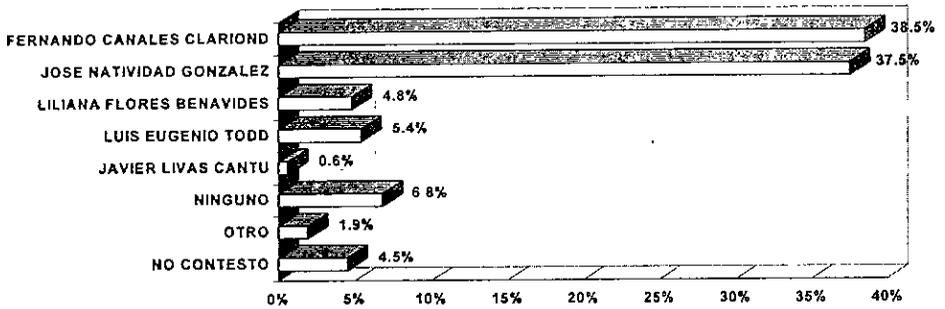


Gráfico 4.2.D Simpatía hacia los candidatos para gobernador del Estado

DE LAS SIGUIENTES PERSONAS, ¿QUIEN LE GUSTARIA QUE FUERA EL PROXIMO GOBERNADOR DE NUEVO LEON?

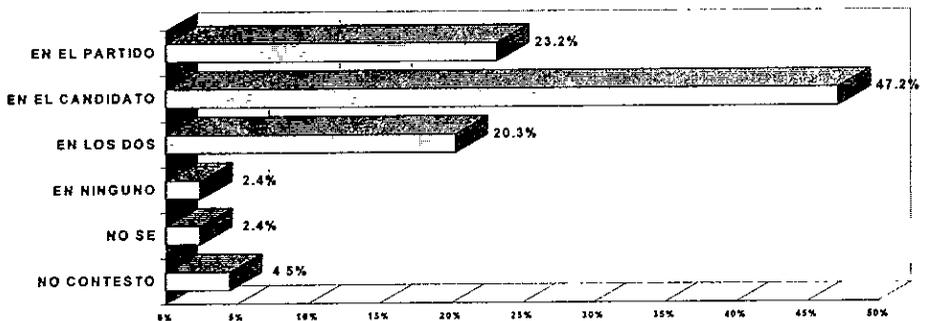
ENCUESTA PREELECTORAL EN LA ELECCION CELEBRADA EL 6 DE JULIO DE 1997



La gráfica 4.2.D indica, a diferencia de la anterior, la preferencia que la población manifiesta hacia los candidatos para gobernador; Canales Clariond del PAN al parecer es el favorito, pero la diferencia con José Natividad del PRI no es significativa, la conclusión es que por su persona ambos tienen posibilidades de ser el próximo gobernador.

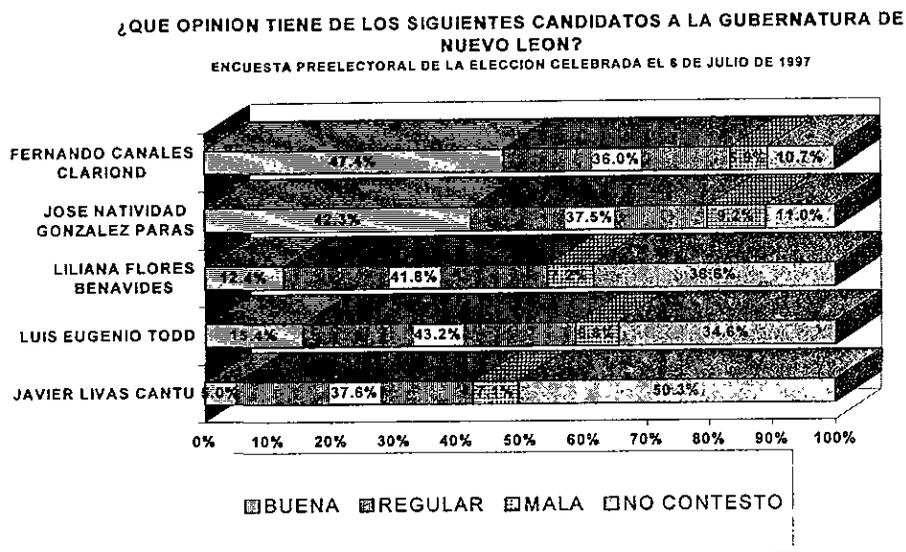
Gráfico 4.2.E Importancia del candidato y el partido en la ciudadanía

AL VOTAR EN QUE SE FIJARIA PRIMERO, ¿EN EL CANDIDATO O EN EL PARTIDO?
ENCUESTA PREELECTORAL EN LA ELECCION A GOBERNADOR DEL ESTADO DE NUEVO LEON CELEBRADA EL 6 DE JULIO DE 1997



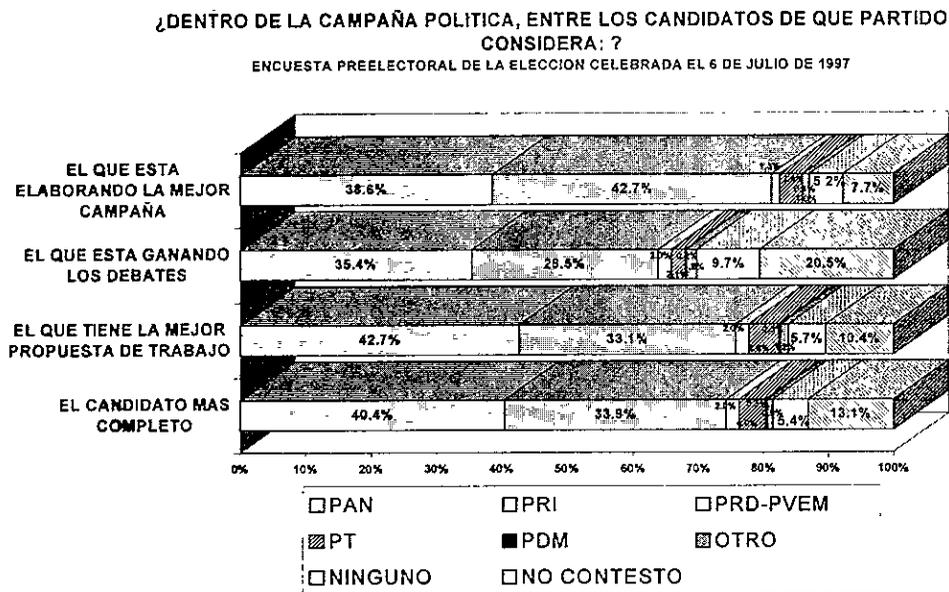
El gráfico 4.2.E indica que para la población es más importante el candidato que el partido político, esta tendencia se ha mantenido en los últimos años no sólo en Nuevo León sino en otros Estados del país como son: Colima, San Luis Potosí, DF, Sonora y Campeche de acuerdo a estudios del Centro de Estudios de Opinión. Lo anterior indica que la población desea un gobernador por encima del partido al que pertenezca.

Gráfico 4.2.F Opinión de la población sobre cada uno de los candidatos



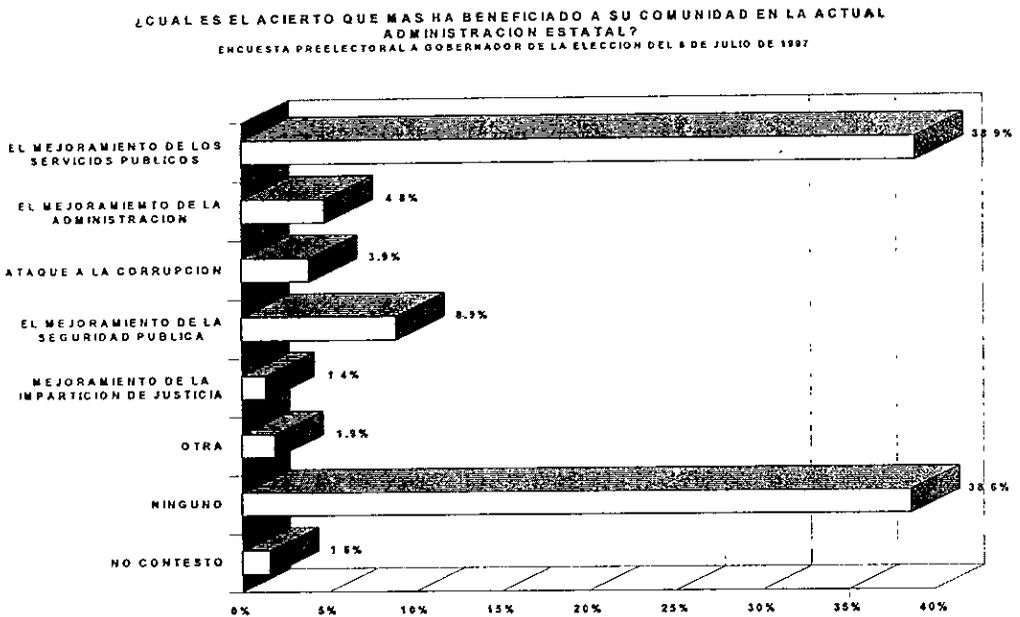
El gráfico 4.2.F indica que tan buena o mala es la imagen que la población manifiesta hacia los candidatos para gobernador en el Estado de Nuevo León; los candidatos del PAN y del PRI son los que, según la población adulta tienen mejor imagen en el Estado, de allí que es probable que la elección a celebrarse el día 6 de julio de 1997 sea una elección cerrada, la gráfica 4.2.E indica que, para la población, es más importante el candidato que el partido, es por eso que este gráfico nos indica que tan competida será la elección, si tomamos en cuenta esta gráfica, se puede ver que el candidato del PAN goza de una mejor opinión y de simpatía para ser gobernador que su más cercano contendiente del PRI; los candidatos de los demás partidos presentan altos índices de no contestó, como factor importante del presente resultado consideramos el poco conocimiento que la población en general tiene de los mismos.

Gráfico 4.2.G Impacto sobre las propuestas de campaña para gobernador de Nuevo León



El gráfico 4.2.G presenta el impacto de las estrategias que los partidos tienen hacia la ciudadanía, se puede ver reflejado en los resultados que la mejor campaña fue del PRI con un 42.7%, contra el 38.6% que mencionan que el PAN fue quien hizo mejor su campaña; las campañas fueron cerradas y competidas, pero el PRI la realizó ligeramente mejor que el PAN. En la siguiente pregunta se relaciona que los debates establecidos los ha ganado el PAN con un 35.4% y que el PRI los ganó con un 28.5% según los respondientes, hasta aquí pareciera ser que la contienda política se encuentra pareja, consideramos que una variable determinante en las elecciones son los debates, sin embargo, el PRI contrarresta la balanza en su contra con una mejor campaña, las siguientes dos preguntas a nuestro juicio son concluyentes, el 42.7% piensa que el PAN tiene la mejor propuesta de trabajo por encima de la del PRI con un 33.1%, además la población piensa que el candidato más completo es el del PAN con un 40.4% contra un 33.9% del PRI. A modo de conclusión, hasta aquí, el PAN tiene ligeras posibilidades de ganar, siendo estas mayores a las del PRI, pues, como se vio en la gráfica 4.2.F, es importante en la decisión de un elector el candidato como individuo, por encima del partido político que representa, la población cree que el candidato del PAN es más completo que el del PRI bajo esta perspectiva.

Gráfico 4.2.H Beneficios hacia la comunidad según encuesta preelectoral en Nuevo León

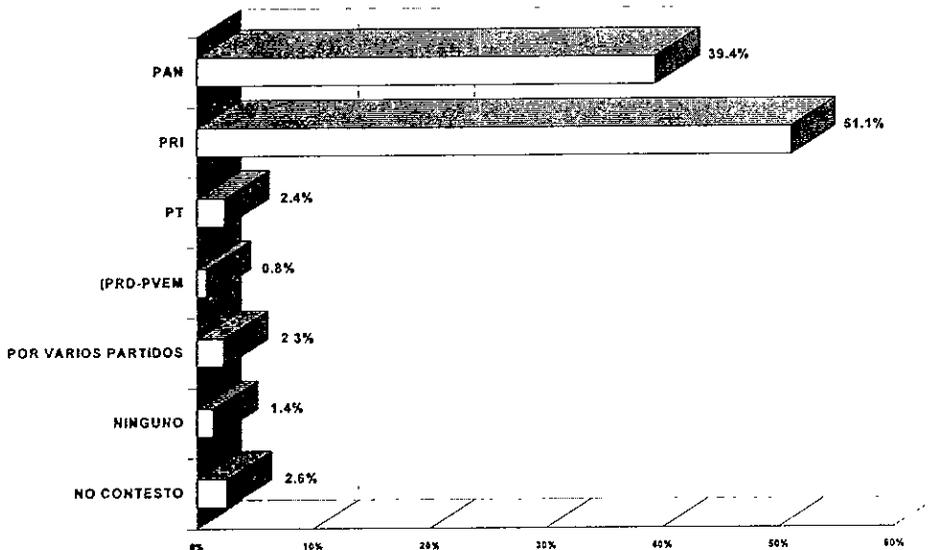


De la actual administración estatal, la población adulta dice que el mejoramiento de los servicios públicos fue el mejor acierto que ha tenido, sin embargo, con el mismo porcentaje se comenta que ningún acierto ha tenido, esta es una variable a considerar seriamente para las aspiraciones del próximo candidato del PRI, ya que ésta es una forma de medir que tan contenta se encuentra la gente con el partido en el poder; si la gente está dispuesta a darle continuidad al proyecto o lo contrario. Al candidato del partido en el poder se le cataloga como la sucesión que le da o no continuidad a las políticas y estrategias económicas y sociales, es por esta razón que el candidato del PRI debería engrandecer en gran medida los logros de la actual administración a la percepción de la población.

El principal oponente al partido en el poder, el PAN, debería dirigir sus esfuerzos a la población que no está contenta con la administración. Los demás reactivos casi no son mencionados, pero uno que es mencionado es que se mejoró la seguridad pública, punto a resaltar en una campaña política.

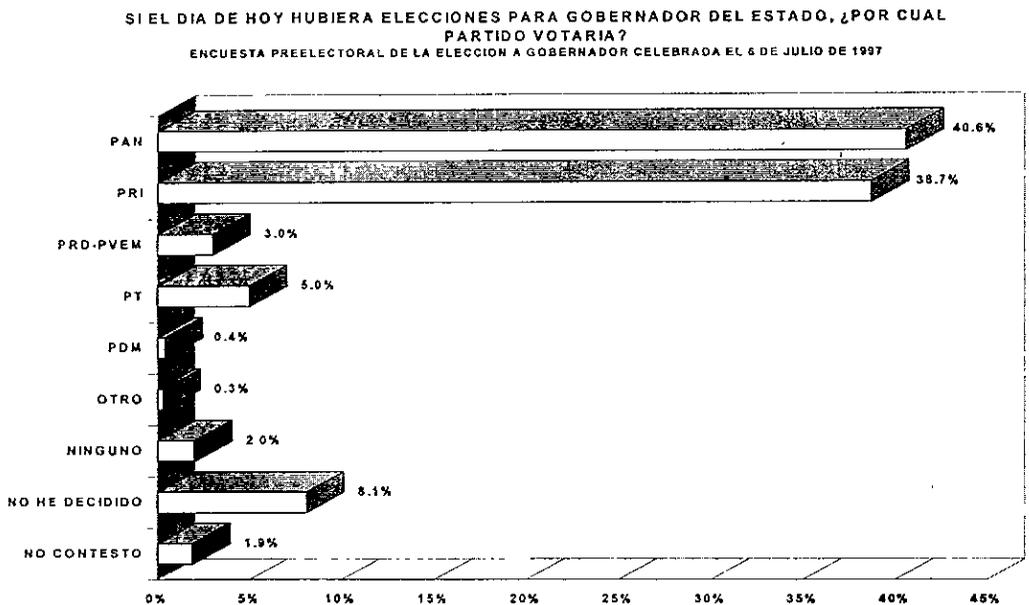
Gráfico 4.2.I *Votación estimada en las pasadas elecciones para gobernador de Nuevo León (1298 casos)*

**¿POR CUAL PARTIDO VOTO EN LAS ELECCIONES ANTERIORES?
ENCUESTA PREELECTORAL DE LA ELECCION CELEBRADA EL 6 DE JULIO
DE 1997
(72.1% DE LOS CIUDADANOS QUE DIJERON VOTAR EN LAS ELECCIONES
PASADAS PARA PRESIDENTE DE MEXICO)**



El gráfico 4.2.I sólo indica la votación de la elección pasada, en la que efectivamente el triunfador fue el PRI, cabe recordar que el gobernador del Estado, antes de la elección del 6 de julio, era de este partido, también se hizo una pregunta que no se puso en gráfico y es si votó en las elecciones pasadas, donde un 72.1% dijeron que sí. Los resultados electorales de la elección para Presidente de México de la pasada elección fueron de un 40.6% para el PAN, un 49.2% para el PRI, ésta gráfica demuestra que los electores de Nuevo León recuerdan por cual partido político votaron las pasadas elecciones. Como dato, en la pasada elección votaron 1'503,737 de un total de 1'899,908 que representaron un 79.15% de participación contra el 72.1% de la encuesta.

Gráfico 4.2.J Intención de voto al momento de la encuesta preelectoral en Nuevo León



Este gráfico 4.2.J es de los más importantes de este estudio preelectoral a nuestro juicio, ya que presenta las tendencias electorales al momento del levantamiento de la encuesta y consideramos difícil que las tendencias en 15 días cambian de rumbo. En apariencia el PAN obtendría el triunfo con un 40.6% de la votación seguido de cerca por el PRI con un 38.7% de la votación, el resto de los partidos políticos no tienen un porcentaje significativo, tal vez el PT con un 5.0% pero sus aspiraciones a obtener la gubernatura del Estado de Nuevo León son prácticamente nulas. La explicación del porque el PT queda por encima de la alianza PRD-PVEM es muy sencilla, cabe recordar la gráfica 4.2.E en el que para la población es más importante el candidato que el partido, en la gráfica 4.2.D se puede ver que el candidato Luis Eugenio Todd (candidato del PT) con un 5.4% tiene mayor simpatía para gobernador del Estado que Liliana Flores Benavides (candidata del PRD-PVEM) con un 4.8%, así se puede explicar la cantidad del porcentaje de votos del PT con respecto al porcentaje de votos del PRD que se considera un partido más fuerte políticamente en todo el país.

Hasta aquí según los resultados del estudio el PAN tanto en candidato como en partido tiene por un pequeño margen posibilidades mayores de ganar la elección, sin embargo, si se consideran los intervalos de confianza a un 95% de la intención de voto del PAN y del PRI dan los siguientes resultados:

Cuadro 4-1 Intervalos de confianza de la encuesta preelectoral de la elección para gobernador de Nuevo León celebrada el 6 de julio de 1997 hacia el PRI y el PAN.

PARTIDO	PARAMETROS		INTERVALO DE CONFIANZA (95%)	
	MEDIA	VARIANZA	INFERIOR	SUPERIOR
PAN	40.60%	0.0646%	35.62%	45.58%
PRI	38.70%	0.0456%	34.51%	42.89%

Al analizar estos resultados muestran que, aunque con una probabilidad muy pequeña, el PAN puede descender hasta 35.62% y el PRI puede ascender hasta 42.89%, así la elección de acuerdo a los resultados puntuales, donde el PAN obtendría el triunfo, se complica por la intersección de los intervalos de confianza, si los intervalos de confianza de dos parámetros se intersectan, se le llama empate técnico, es decir, que no se está en condiciones de mencionar cual es el partido que en el momento del levantamiento tiene la mayoría relativa con una confianza del 95%, asimismo recordemos que las posibilidades de que el PAN baje a su nivel más bajo y el PRI suba al más alto son poco probables, se debe confiar en los resultados puntuales.

Otro punto importante de mencionar es que un 8.1% de los encuestados se encuentran indecisos, la diferencia entre el PAN y el PRI es de 1.9% puntos sobre los estimadores puntuales, de allí la pregunta obligada, ¿Los indecisos hacia que partidos y de que manera se distribuirán?, una declaración importante, o negativa de algún candidato, o de los miembros importantes de los partidos políticos juegan un papel determinante sobre todo en los últimos días anteriores al día de la elección.

Los partidos políticos, de tener conocimiento de los resultados estarían realizando una campaña dirigida hacia los indecisos para que se decidan por determinado partido, los indecisos por consecuencia juegan un papel importante en la encuesta preelectoral, bajo este análisis no se está en condiciones de decir quien será el próximo gobernador del Estado de Nuevo León con total certeza, pero si confiamos en los resultados puntuales el ganador sería el candidato del PAN.

Gráfico 4.2 K Por que votarían por el PAN en Nuevo León (731 casos)

¿POR QUE VOTARIA POR ESE PARTIDO PARA GOBERNADOR DEL ESTADO?
ENCUESTA PREELECTORAL DE LA ELECCION DEL 6 DE JULIO DE 1997
(40.6% DE LOS CIUDADANOS QUE VOTARIAN POR EL PAN)

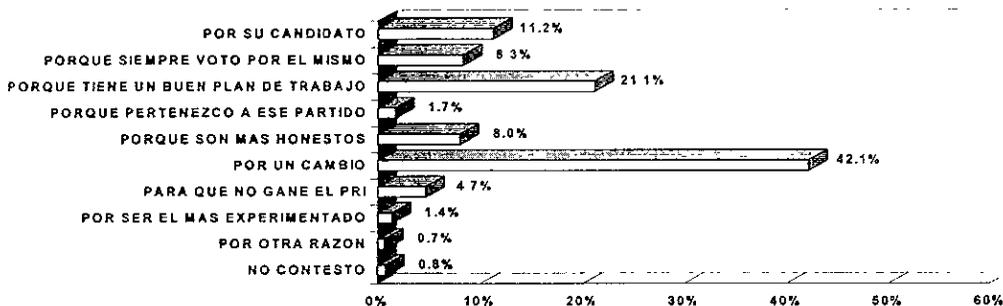
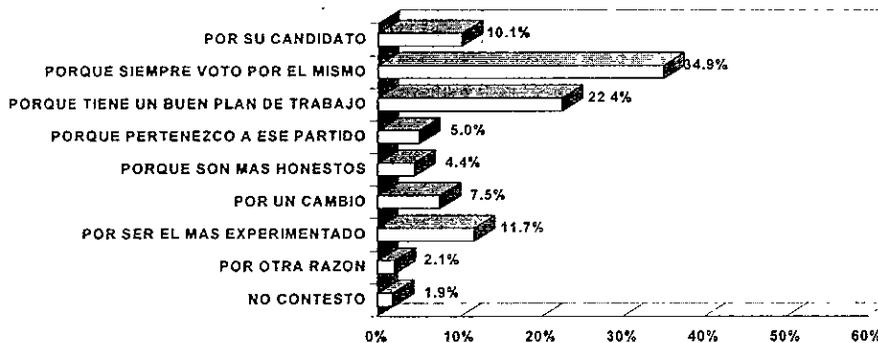


Gráfico 4.2.L Por que votarían por el PRI en Nuevo León (697 casos)

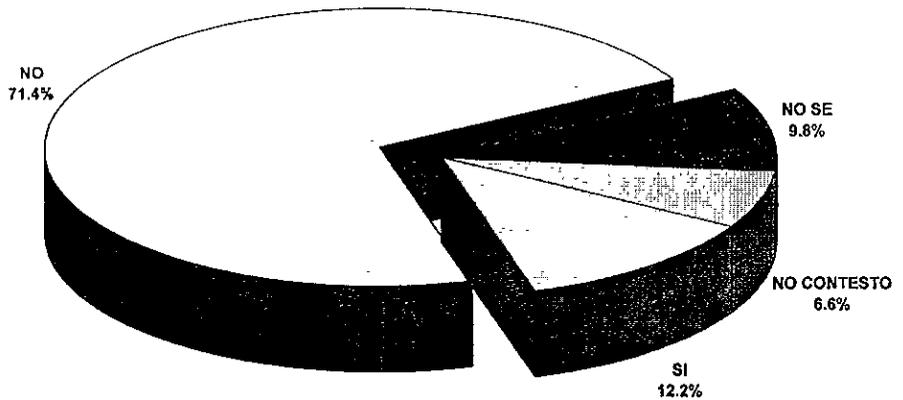
¿POR QUE VOTARIA POR ESE PARTIDO PARA GOBERNADOR DEL ESTADO?
ENCUESTA PREELECTORAL EN LA ELECCION DEL 6 DE JULIO DE 1997
(38.7% DE LOS CIUDADANOS QUE VOTARIAN POR EL PRI)



Estas gráficas (4.2.K y 4.2.L) muestran que la mayoría de los que votarían por el PAN es por que buscan un cambio, en tanto que la mayoría de los que votarían por el PRI dicen que votan siempre por el mismo, a este tipo de voto se le llama voto duro que es un voto característico del PRI, sin embargo, una buena parte mencionan que les gusta el plan de trabajo.

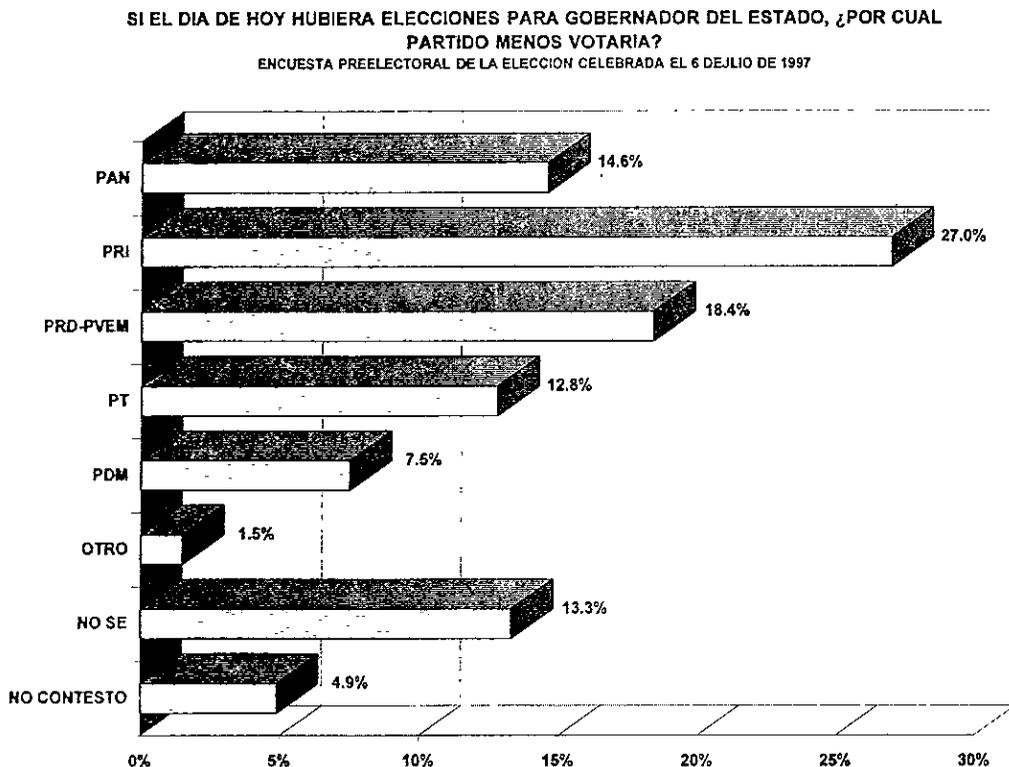
Gráfico 4.2.M Posibilidad de cambio de decisión en la elección de Nuevo León

¿CONSIDERA PROBABLE CAMBIAR SU DECISION DE VOTO DE AQUI AL DIA DE LAS ELECCIONES A GOBERNADOR DEL ESTADO DE NUEVO LEON?
ENCUESTA PREELECTORAL DE LA ELECCION CELEBRADA EL 6 DE JULIO DE 1997



El gráfico 4.2.M indica la posibilidad de que los electores cambien su decisión el día de la elección, el 71.4% menciona que la decisión está tomada y ya decidieron por quien van a votar, el 12.2% dicen que es probable que cambien de decisión, un 9.8% menciona que no sabe si cambiará de decisión, mientras que sólo un 6.6% no contestaron a este reactivo.

Gráfico 4.2.N Animadversión hacia los partidos políticos de la elección en Nuevo León



La gráfica 4.2.N proporciona la animadversión que la población tiene hacia los partidos políticos, es decir, el rechazo hacia éstos.

El PRI tiene una animadversión mayor de la población con un 27.0% en el Estado de Nuevo León, es el partido político que mayor porcentaje de rechazo tiene, en tanto la alianza PRD-PVEM con una animadversión de 18.4% es mayor que la que tiene el PAN con un 14.6%; es un gráfico más que indica que el PAN será el partido político con mayores posibilidades de ganar en las elecciones del 6 de julio de 1997 a la gubernatura del Estado de Nuevo León, ya que recibe el menor rechazo, por lo tanto y en contraste se manifiesta la mayor aceptación.

Los intervalos de confianza que se presentan en el cuadro 4-2 sobre la animadversión nos indican que el rechazo máximo hacia el PAN fue de 16.53%, el rechazo mínimo hacia el PRI fue de un 23.63%, desde luego que esto es un indicador positivo hacia el PAN que tiene una animadversión más pequeña incluso que el PRD y similar al PT, ambos son partidos que no alcanzaron alta votación a pesar de haber sido de los partidos políticos más importantes en la contienda electoral.

Cuadro 4-2 Intervalos de confianza de la animadversión de la encuesta preelectoral de la elección para gobernador de Nuevo León celebrada el 6 de julio de 1997.

PARTIDO	PARAMETROS		INTERVALO DE CONFIANZA (95%)	
	MEDIA	VARIANZA	INFERIOR	SUPERIOR
PAN	14.6%	0.0097%	12.67%	16.53%
PRI	27.0%	0.0296%	23.63%	30.37%
PRD-PVEM	18.4%	0.0213%	15.54%	21.26%
PT	12.8%	0.0135%	10.52%	15.08%

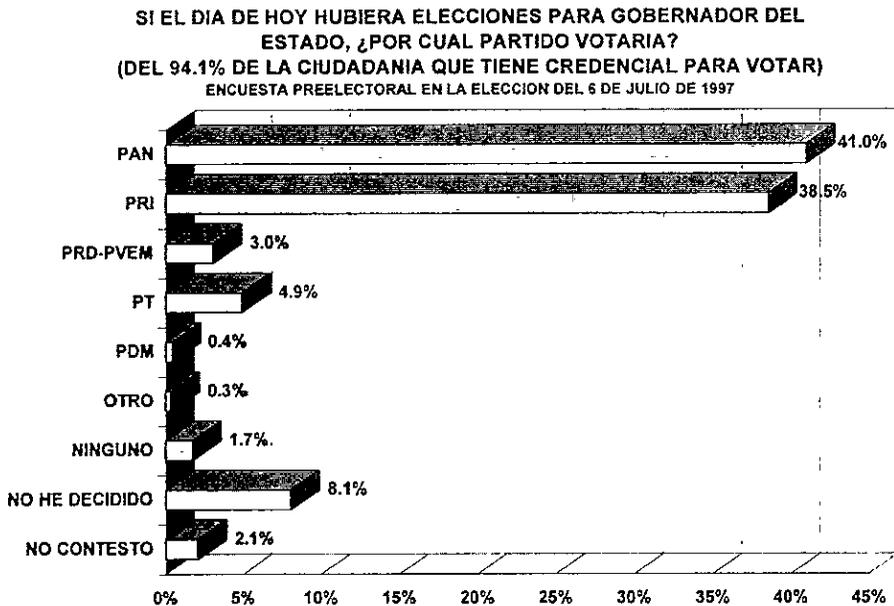
Los intervalos del cuadro 4-2 son a un 95% de confianza.

4.2.2 CRUCES DE VARIABLES

Un cruce de variables es cuando se examina una pregunta con cierta característica de interés en relación a otra, ésta puede ser demográfica o de otra pregunta, un ejemplo sería observar la intención de voto de los que tienen credencial para votar con fotografía.

En el gráfico 4.2.O se puede ver la intención del voto de las personas que sí tienen credencial para votar con fotografía, estas personas son las que están en condiciones de emitir su voto. El siguiente gráfico puede decir la intención del voto con mayor certeza que el gráfico 4.2.J, debido a que sólo toma la intención del voto para aquellos que están en condiciones de acudir a votar.

Gráfico 4.2.O Intención de voto de los que tienen credencial en el Estado de Nuevo León (1694 casos)

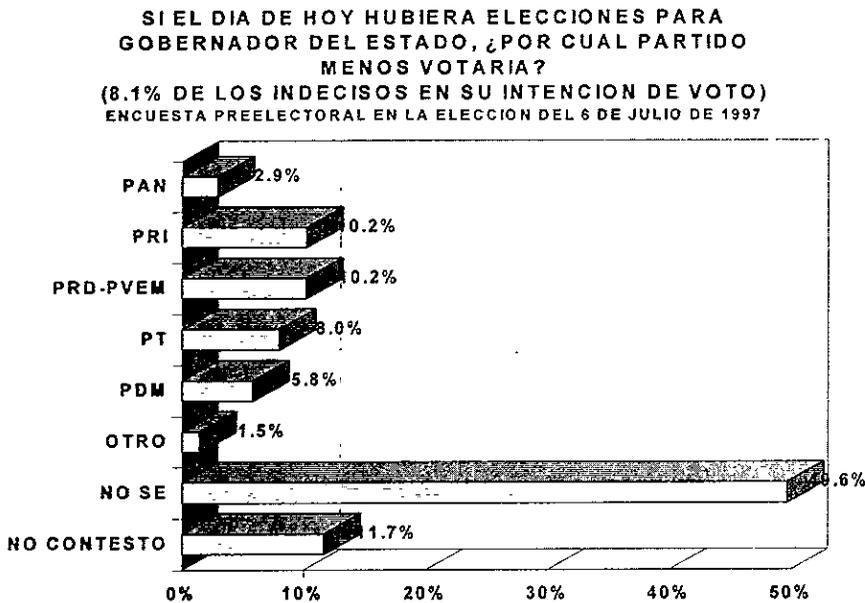


Este gráfico (4.2.O) no difiere mucho del 4.2.J, ya que visualizando este resultado se puede observar que la tendencia del PRI y del PAN se mantienen de forma igual.

Un cruce de variables que consideramos importante para el análisis de las tendencias de la intención del voto son los indecisos, si bien son personas que no saben por alguna razón por quien van a votar, es probable que si sepan por quien no votarán. De esta manera dentro de los indecisos se puede ver quien tiene menor animadversión a determinado partido político y es probable que éstos le favorezcan.

Según el gráfico 4.2.P el PAN tiene menor animadversión que el PRI de allí que se puede suponer que los indecisos favorecerán al PAN, entonces tenemos una razón más para suponer que existen posibilidades de que el PRI pierda la gubernatura del Estado de Nuevo León en las próximas elecciones.

Gráfico 4.2.P Animadversión de los indecisos en el Estado de Nuevo León (146 casos)



A pesar que de manera relativa el PAN tiene menor animadversión que el PRI su principal adversario para la elección del 6 de julio, se puede ver que el 49.6% no saben por cual partido político no votarían. Por consiguiente no se puede decir que los indecisos favorecerán al PAN el día de las elecciones.

En los siguientes gráficos 4.2.Q y 4.2.R se puede ver el cruce de la intención de voto por género.

Gráfico 4.2.Q Intención de voto de los hombres en el Estado de Nuevo León (892 casos)

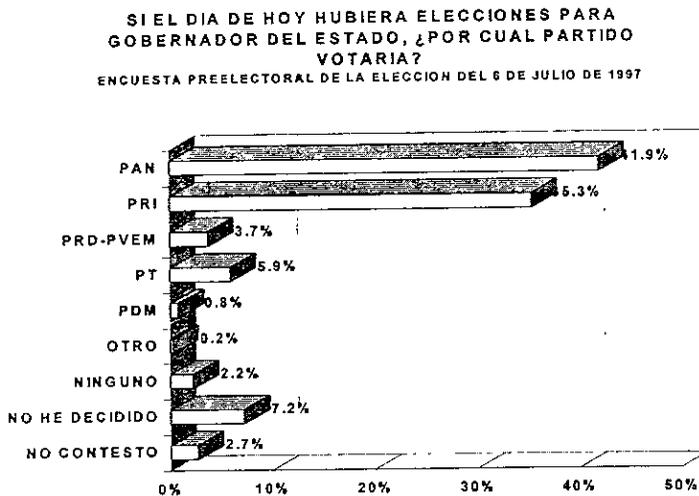
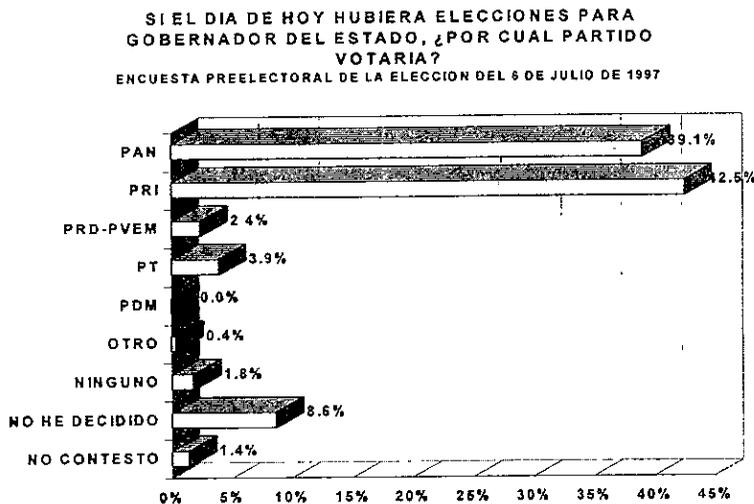


Gráfico 4.2.R Intención de voto de las mujeres en Nuevo León (908 casos)



Se puede ver que el voto de los hombres favorece al PAN con un 41.9% contra un 35.3% del PRI; mientras que el voto de las mujeres favorece al PRI con un 42.5% contra un 39.1% al PAN.

Cuadro 4-3: Intención de voto en porcentajes según edad en la encuesta preelectoral de la elección celebrada el 6 de julio de 1997 para gobernador del Estado de Nuevo León

EDAD	PAN	PRI	PRD	PT	OTRO	NINGUNO	INDECISO	NC	TOTAL	CASOS
18-24	44.09	38.41	3.86	3.18	0.9	1.59	5.91	2.06	100.00	466
25-34	44.40	35.85	2.44	4.68	0.61	3.26	7.74	1.02	100.00	520
35-44	39.23	38.94	2.95	5.90	0.88	1.77	9.14	1.19	100.00	359
45 y >	34.19	42.09	2.79	6.05	0.47	1.16	9.30	3.95	100.00	455

Se puede ver que los jóvenes de 18 a 24 años votan en su mayoría por el PAN, el bloque que más beneficia al PAN es el de 25 a 34 años, el grupo de 35 a 44 también votan por el PAN aunque la diferencia con el PRI es moderada, por último podemos ver que el grupo de 45 y más favorece ampliamente al PRI. Lo que nos indica que la juventud está a favor del cambio de partido político en el Gobierno.

El siguiente cuadro es un cruce de la intención de voto con la animadversión del voto.

Cuadro 4-4 Cruce: intención de voto vs animadversión en porcentajes en la encuesta preelectoral de las elecciones del 6 de julio de 1997 para gobernador de Nuevo León

Intención de voto

PARTIDO	PAN	PRI	PRD-PVEM	PT	PDM
PAN		31.62	14.28	15.66	16.67
PRI	49.48		55.10	54.21	50.01
PRD-PVEM	18.91	23.95		12.04	0.0
PT	13.15	16.90	8.16		0.0
PDM	7.82	8.76	10.21	7.23	
OTRO	1.33	1.09	4.09	3.61	0.0
NO SE	7.38	13.92	4.09	3.61	0.0
NC	1.93	3.76	4.08	3.63	33.32
TOTAL	100.0	100.0	100.0	100.0	100.0
CASOS	734	697	59	90	7

Animadversión

Este cuadro 4-4, muestra que aquellas personas que votan por el PAN no lo harían por el PRI con un 49.48%, las que votan por el PRI no votarían por el PAN como primera opción con un 31.62%, aquellos que votarían por el PRD-PVEM la mayoría (un 55.10%) no lo haría hacia el PRI. De manera similar ocurre con el PT y el PDM.

4.2.3 COMPARATIVOS DE LA ENCUESTA PREELECTORAL

Las cuotas se definieron en el apartado 2.7, sin embargo, una condición que se pide es que los porcentajes de éstas sean iguales a la realidad, la información que se considera oficial se obtuvo a partir del Censo del 95 publicado por el INEGI.

Cuadro 4-5 Porcentajes de género mayor de 18 años según Censo95 del Inegi y encuesta preelectoral de la elección para gobernador celebrada el 6 de julio de 1997 en el Estado de Nuevo León

	HOMBRES	MUJERES
ENCUESTA	49.53%	50.47%
REAL	49.20%	50.80%
DIFERENCIA	0.33%	-0.33%

Se puede ver que la diferencia de los porcentajes de cuotas según género en la encuesta comparada con la real es negligible, cabe recordar que la edad fue una cuota dada como requisito hacia los encuestadores.

Cuadro 4-6. Porcentajes de grupos de edad mayor de 18 años según Censo95 del Inegi y encuesta preelectoral de la elección para gobernador celebrada el 6 de julio de 1997 en Nuevo León

EDAD	POB > 18	ENCUESTA	DIF
18-24	27.37%	25.88%	1.49%
25-34	27.51%	28.88%	-1.37%
35-44	18.13%	19.95%	-1.82%
45 y >	26.98%	25.30%	1.69%

La otra cuota aplicada en la muestra del Estado de Nuevo León fue grupos de edad, las diferencias son pequeñas como puede verse en el cuadro 4-6, estas cuotas de alguna manera no representan sorpresa alguna, pues, a los encuestadores se les indicó que aplicarían esto en muestra. Existe otra variable que es nivel de instrucción que se codificó para el análisis de posibles sesgos en la muestra, ver cuadro 4-7, esta variable no se les indicó cantidad alguna a los encuestadores, así se está en posibilidades de valorar la muestra con respecto a los resultados que publica el INEGI en el Censo95, las diferencias no son amplias. Al menos en las variables analizadas se puede decir que la muestra está balanceada.

Cuadro 4-7 Porcentajes de grupos según nivel de instrucción de mayores de 18 años según Inegi y encuesta preelectoral de la elección para gobernador celebrada el 6 de julio de 1997 en el Estado de Nuevo León

NIVEL DE INSTRUCCION	PREELECTORAL	OFICIAL	DIF
SIN ESTUDIOS	4.37%	6.50%	-2.13%
PRIMARIA	32.36%	35.89%	-3.53%
SECUNDARIA	27.62%	23.23%	4.39%
BACHILLERATO Y TEC.	22.41%	21.06%	1.35%
LICENCIATURA	12.70%	11.78%	0.92%
POSTGRADO	0.54%	1.54%	-1.00%

4.2.4 VARIANZA DE INTENCION Y ANIMADVERSION DEL VOTO

En este apartado se presentan las varianzas de la intención y animadversión del voto en el Estado de Nuevo León. Las varianzas corresponden a los principales partidos políticos y se dan por estrato y en el ámbito estatal.

Cuadro 4-8 Varianzas de la intención del voto de la encuesta preelectoral por estrato de la elección para gobernador celebrada el 6 de julio de 1997 en el Estado de Nuevo León

N	ESTRATO	PAN	PRI	PRD-PVEM	PT
1	MONTERREY	0.0741%	0.0692%	0.0053%	0.0123%
2	GUADALUPE	0.2103%	0.0981%	0.0086%	0.0155%
3	SAN NICOLAS DE LOS GARZA	0.1715%	0.0949%	0.0165%	0.0138%
4	APODACA Y ESC	0.1546%	0.3522%	0.0154%	0.0419%
5	SAN PEDRO GARZA GARCIA	0.4996%	0.2539%	0.2946%	0.0555%
6	SANTA CATARINA	0.8749%	0.5522%	0.0292%	0.0418%
7	RESTO	1.3785%	0.9307%	0.0025%	0.0696%
T	TOTAL	0.0646%	0.0456%	0.0019%	0.0047%

El estrato que mayor varianza tiene hacia el PAN es el resto del Estado, es decir, el estrato 7, es el estrato que abarca 44 municipios de los cuales se tomaron 10 para la muestra y la forma de pensar para habitantes de un mismo municipio generalmente es más parecida que personas de distintos municipios; lo mismo ocurre con la varianza hacia el PRI. Para el caso del estrato 1, es decir, Monterrey, la varianza es la más pequeña de todos los estratos, pero cabe recordar que en el estrato siete se utilizó un modelo trietápico y en los demás estratos un bietápico.

Para el caso de la animadversión del voto se puede observar el cuadro 4-9 a continuación.

Cuadro 4-9 Varianzas de la animadversión del voto de la encuesta preelectoral por estrato de la elección para gobernador celebrada el 6 de julio de 1997 en el Estado de Nuevo León

N	ESTRATO	PAN	PRI	PRD-PVEM	PT
1	MONTERREY	0.0266%	0.0656%	0.0235%	0.0212%
2	GUADALUPE	0.0778%	0.1744%	0.0586%	0.0548%
3	SAN NICOLAS DE LOS GARZA	0.0389%	0.0744%	0.1079%	0.0813%
4	APODACA Y ESC	0.1150%	0.1156%	0.2116%	0.0944%
5	SAN PEDRO GARZA GARCIA	0.0623%	0.5002%	0.2697%	0.3679%
6	SANTA CATARINA	0.0903%	0.5785%	0.2427%	0.0879%
7	RESTO	0.0800%	0.4015%	0.3984%	0.2096%
T	TOTAL	0.0097%	0.0296%	0.0213%	0.0135%

4.2.5 EFECTO DE DISEÑO DE LA ENCUESTA PREELECTORAL

El efecto de diseño presentado corresponde a la principal variable de la encuesta que fue la intención del voto. Las varianzas del partido político mayoritario en esta elección, es decir el PAN en relación con el diseño muestral elaborado para esta encuesta y bajo un modelo de muestreo aleatorio simple son las siguientes:

$$V(\hat{p}_{PAN}) = 0.0646\%, \quad (4.2.1)$$

$$V_{max}(\hat{p}_{PAN}) = 0.0349\%, \quad (4.2.2)$$

dividiendo (4.2.1) entre (4.2.2) nos queda el efecto de diseño:

$$Deff = 1.8534, \quad (4.2.3)$$

utilizando (4.2.3) para el tamaño de muestra a un 95% de confianza y un error de $\pm 5\%$ de error absoluto nos queda:

$$n = n_{max} Deff = 371 * 1.8534 \cong 687. \quad (4.2.4)$$

4.3. ENCUESTA DE SALIDA O POSTVOTO

4.3.1 RESULTADOS DE LA ENCUESTA DE SALIDA

En esta sección se pueden visualizar los resultados del primer corte parcial, del segundo corte parcial y su acumulado, del tercer corte tanto parcial como acumulado. Al tercer corte acumulado se le llama encuesta de salida o postvoto. Todos los resultados se pueden visualizar en un solo gráfico salvo el tercer corte acumulado que se definió como la encuesta de salida y engloba toda la jornada electoral.

Los resultados de la encuesta de salida se pueden visualizar por estrato para conocer el comportamiento que tiene cada uno de ellos. Además se observa el gráfico comparativo de cada uno de los tres cortes parciales. Por último, se puede visualizar una curva de la evolución del voto en los tres cortes acumulados.

En el gráfico 4.3.A se pueden ver los resultados que hasta las 10:30 hrs del día 6 de julio de 1997 se obtuvieron en la encuesta de salida del primer corte, donde se observa que el PAN obtuvo un 38.1% y el PRI obtuvo 32.4%, de esta manera a las 10:30 hrs el 1er corte determina una ventaja de 5.70% del PAN hacia el PRI. Existe un alto índice de no contestó con un 23.1%, la pregunta es por quien votaron los que no contestaron, la realidad es que no se puede saber. Si suponemos que la distribución del voto de los que no contestaron se hizo de manera proporcional a la intención de voto de los que sí contestaron en el primer corte, es decir, se extrapolaron los que no contestaron con los que sí lo hicieron y así se obtiene la gráfica 4.3.B

El gráfico 4.3.A indica que el PAN en el segundo corte parcial, al igual que en el primero se mantiene como un fuerte candidato a ocupar la Gubernatura del estado de Nuevo León, sin embargo, se puede observar que siguió existiendo un alto grado de no contestó con un 22.0%, que no se sabe en que proporciones se distribuye hacia los partidos políticos, el triunfo aparente en el segundo corte parcial es para el PAN; este triunfo parcial sólo abarca de las 11:00 hrs a las 14:30 hrs. Si se supone al igual que el primer corte, que la distribución de los no contestó es de manera proporcional sobre los que sí lo hicieron, se construye el gráfico 4.3.B donde se observa la intención de voto de cada uno de los partidos políticos.

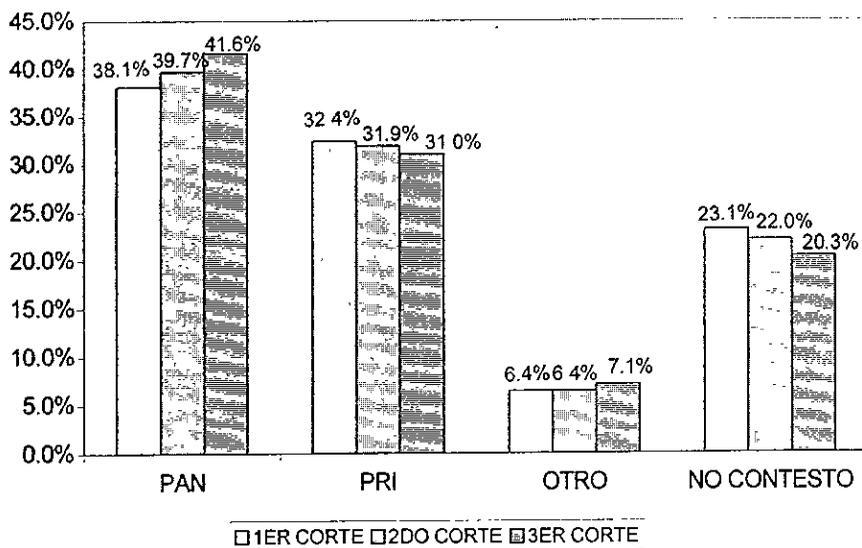
En el tercer corte parcial que abarca y menciona las preferencias electorales de las 15:00 hrs a las 18:00 hrs los resultados se pueden ver en el gráfico 4.3.A.

El tercer corte es el que más favorece al PAN sobre el PRI, las preferencias del PRI como se puede ver son del 31.0%, mientras que las del PAN son del 41.6%, la diferencia es de 10.6%, sin embargo, se puede ver que prevaleció un alto índice del no contestó con un 20.3% del que no se sabe en que porcentajes se distribuyen hacia los partidos políticos. Si una vez más se supone que el no contestó se distribuye de manera proporcional en los que sí contestaron su preferencia electoral se construye el gráfico 4.3.B.

A continuación se presenta el cuadro comparativo del primer corte, segundo corte y tercer corte parciales, para ver el comportamiento de la votación en los diferentes horarios según la encuesta de salida.

Gráfico 4.3.A Postvoto en los cortes parciales con no contestó en el estado de Nuevo León

ENCUESTA POSTVOTO EN LA ELECCION A GOBERNADOR DEL 6 DE JULIO DE 1998

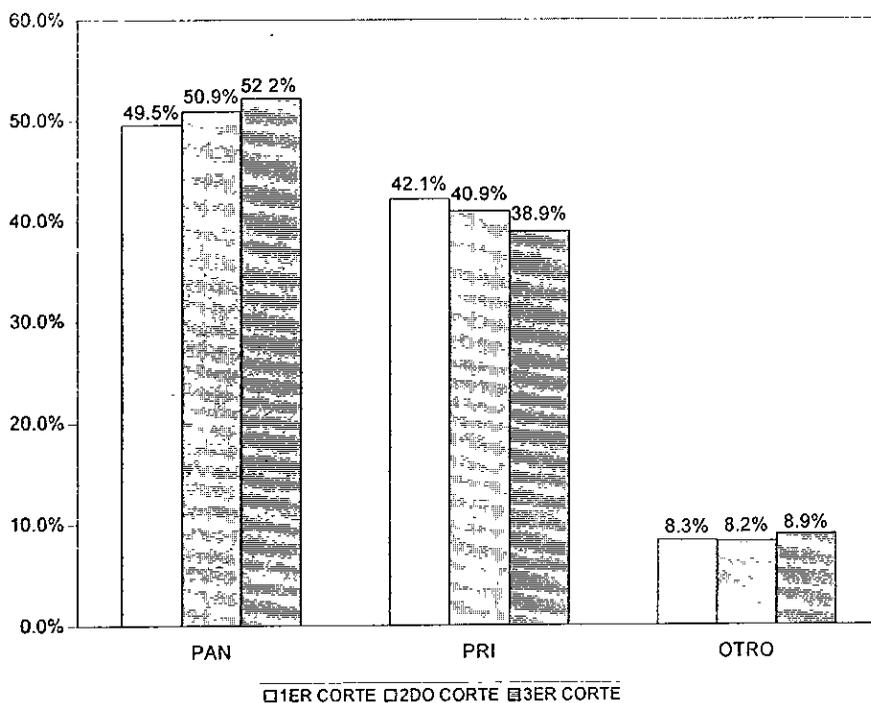


Se puede ver que el voto fuerte del PRI se da por el horario matutino que abarca desde las 8:00 hrs hasta las 10:30 hrs y tiende a bajar la tendencia a través de cada uno de los cortes subsecuentes. El voto fuerte del PAN está en el corte último que abarca desde las 15:00 hrs hasta las 18:00 hrs y su punto más débil es en el horario matutino, los que votan por otros partidos tienden a hacerlo con mayor fuerza en el último corte, sin embargo, las diferencias de cada partido con respecto a cada corte son muy pequeñas y el cambio sufrido fue insignificante para fines prácticos, la no respuesta se mantuvo estable al igual que las demás opciones.

También aquí se presenta la gráfica sin “no contestó” bajo el simple supuesto que los que no contestaron se distribuyen de manera proporcional sobre los que sí respondieron por quién votaron.

Gráfico 4.3.B Postvoto en los cortes parciales sin no contestó en el estado de Nuevo León

ENCUESTA POSTVOTO DE LA ELECCION A GOBERNADOR DEL 6 DE JULIO DE 1997

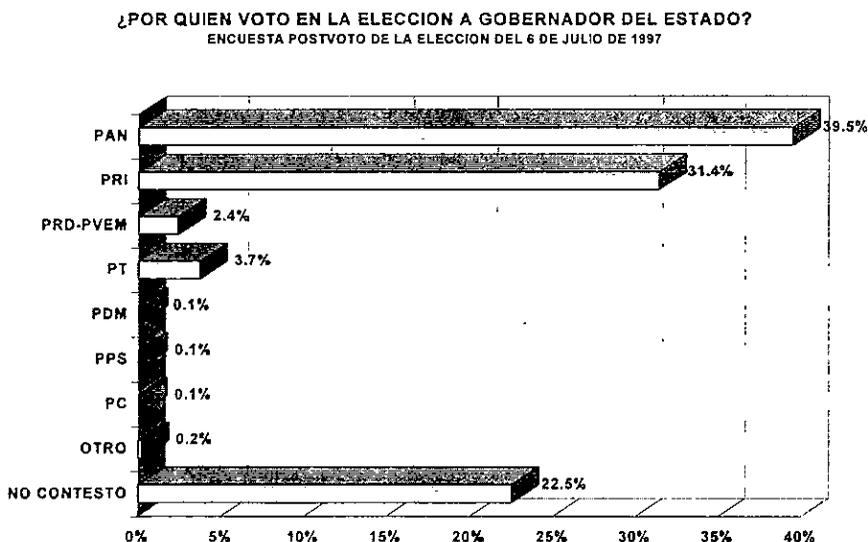


Si sólo se tuviese a nuestra disposición el gráfico 4.3.B, se estaría en condiciones de mencionar el ganador de la Gubernatura del estado de Nuevo León.

Los gráficos 4.3.E y 4.3.F indican la intención de voto del electorado que acudió a votar, pero a diferencia de un corte parcial, este corte refleja la votación total hasta el momento del cierre del mismo, es decir, hasta las 14:30hrs si es el segundo corte acumulado y hasta las 18:00 si es la encuesta de salida global, los resultados aquí presentados son con los que no contestaron por que partido político votaron y el otro resultado se hizo distribuyendo los que no contestaron de manera proporcional sobre los que sí mencionaron por quién votaron, los que no contestan son los electores entrevistados que se niegan a mencionar por cual partido votó.

El gráfico 4.3.E indica que las tendencias hasta el momento del cierre del segundo corte el PAN tiene ventaja sobre el PRI, por consiguiente se estaría en condiciones de afirmar que el PRI perdería la Gubernatura en el estado de Nuevo León, sin embargo, existen una vez más los que no contestaron y no se sabe la distribución de éstos hacia los partidos políticos. El gráfico 4.3.C indica el resultado de la encuesta de salida con no contestó.

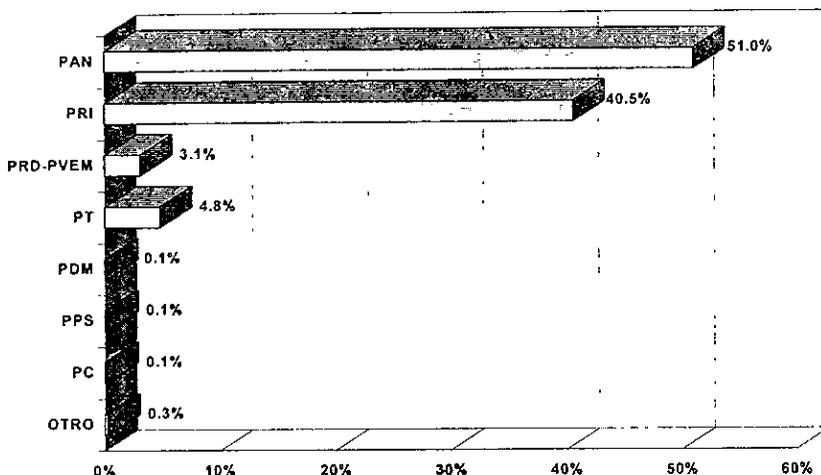
Gráfico 4.3.C Postvoto de la elección en Nuevo León con no contestó



El gráfico 4.3.D sólo es para visualizar como los que no contestaron en el gráfico 4.3.C se distribuyeron de manera proporcional sobre los que sí mencionaron por cual partido político votaron.

Gráfico 4.3.D Postvoto de la elección en Nuevo León sin no contestó

¿POR QUIEN VOTO EN LA ELECCION A GOBERNADOR DEL ESTADO?
ENCUESTA POSTVOTO DE LA ELECCION DEL 6 DE JULIO DE 1997



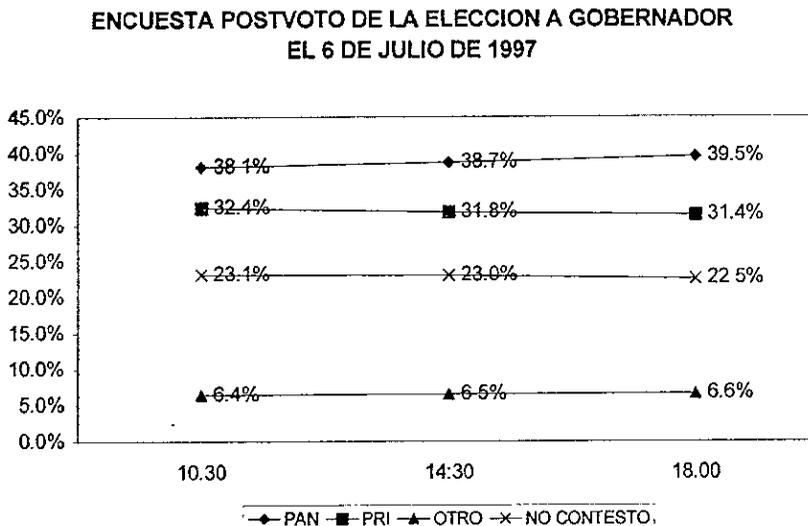
Según el gráfico 4.3.C, sin disponer de mayor información, pone como triunfador de las elecciones a la Gubernatura del estado de Nuevo León al PAN y el PRI pierde la elección y como consecuencia la Gubernatura. Los resultados por estrato en postvoto se pueden ver en el cuadro (4-10):

Cuadro 4-10: Resultados de la encuesta de salida de la elección del 6 de julio de 1997 para gobernador de Nuevo León según estrato

N	ESTRATO	PAN	PRI	PRD-PVEM	PT	OTRO	NC	TOTAL
1	MONTERREY	38.2%	29.8%	2.8%	4.5%	0.5%	24.2%	100.0%
2	GUADALUPE	44.0%	29.5%	3.2%	2.6%	0.5%	20.2%	100.0%
3	SAN NICOLAS	46.2%	26.0%	2.8%	3.4%	0.6%	21.0%	100.0%
4	APODACA Y G. ESCOBEDO	30.1%	37.0%	1.6%	5.2%	0.0%	26.1%	100.0%
5	SAN PEDRO G.	50.5%	19.7%	1.3%	1.3%	0.0%	27.2%	100.0%
6	STA CATARINA	35.5%	35.0%	2.4%	3.3%	0.2%	23.6%	100.0%
7	OTROS	29.9%	39.1%	1.5%	3.4%	0.6%	25.5%	100.0%
T	TOTAL	39.5%	31.4%	2.4%	3.7%	0.5%	22.5%	100.0%

Un gráfico interesante que indica la evolución del voto de la encuesta de salida o postvoto es el 4.3.E con “no contestó”:

Gráfico 4.3.E Evolución del voto en Postvoto a través de los cortes con no contestó en el estado de Nuevo León

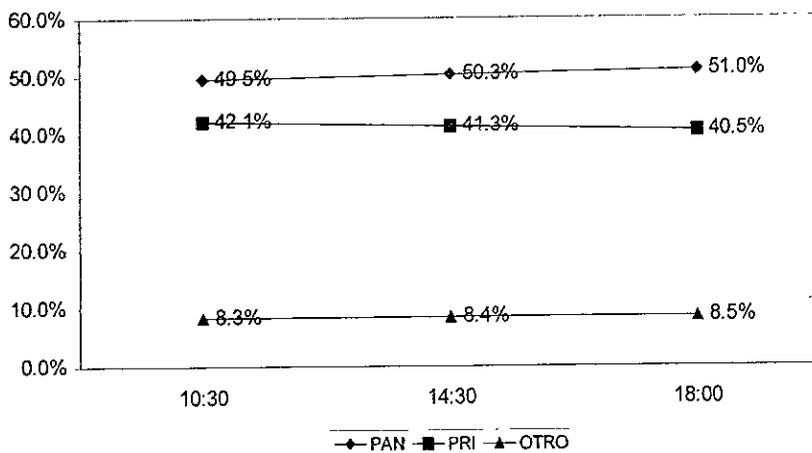


Se observa que a lo largo de la encuesta de salida en el gráfico 4.3.E, la opción “otro” y “no contestó” tienen una pendiente de casi cero, el PRI muestra un pendiente ligeramente negativa y el PAN una pendiente ligeramente positiva a lo largo de la jornada, el decrecimiento y crecimiento respectivamente es casi nulo, pero se puede visualizar que el voto fuerte hacia el PRI se genera por el turno matutino y tiende a decrecer, mientras que el voto hacia el PAN tiende a crecer.

El gráfico 4.3.F se hace bajo la suposición que los que “no contestaron” se distribuyen de manera proporcional bajo los que sí lo hicieron.

Gráfico 4.3.F Evolución del voto en Postvoto a través de los cortes sin no contestó en el estado de Nuevo León

ENCUESTA POSTVOTO EN LA ELECCION A GOBERNADOR
DEL 6 DE JULIO DE 1997



4.3.2 VARIANZA E INTERVALO DE CONFIANZA

En este apartado se presenta la varianza por estrato y en el ámbito estatal de la encuesta de salida, sin embargo, para estimadores de razón con sesgo en muestras de tamaño moderado no sabemos si es válida. Como regla de operación, ésta se puede usar si el tamaño de muestra excede a 30 y que el coeficiente de variación del promedio total de votos sea menor al 10% Cochran (1984, págs 198-199).

Cuadro 4-11: Varianzas de la encuesta de salida de la elección del 6 de julio de 1997 para gobernador de Nuevo León según estrato.

ESTRATO	PAN	PRI	PRD- PVEM	PT	CV (VOTOS)
1	0.0544%	0.0434%	0.0029%	0.0137%	6.1%
2	0.0803%	0.0813%	0.0059%	0.0080%	7.5%
3	0.0801%	0.0735%	0.0069%	0.0111%	7.2%
4	0.1636%	0.1692%	0.0096%	0.0210%	19.1%
5	0.4195%	0.3650%	0.0134%	0.0213%	9.7%
6	0.2374%	0.2430%	0.0217%	0.0265%	16.6%
7	1.1354%	0.6105%	0.0008%	0.0237%	13.7%
TOTAL	0.0473%	0.0299%	0.0008%	0.0030%	5.8%

Los intervalos de confianza a un 95% hacia el PRI y el PAN que son los partidos políticos mayoritarios, se presentan en el cuadro siguiente:

Cuadro 4-12 Intervalos de confianza de la encuesta de salida de la elección para gobernador de Nuevo León celebrada el 6 de julio de 1997 hacia el PRI y el PAN.

PARTIDO	PARAMETROS		INTERVALO DE CONFIANZA (95%)	
	MEDIA	VARIANZA	INFERIOR	SUPERIOR
PAN	51.0%	0.0473%	46.7%	55.3%
PRI	40.5%	0.0299%	37.1%	43.9%

Si se analizan los resultados, con cierta probabilidad el PAN puede tener un descenso hasta 46.7% y el PRI puede tener un ascenso hasta 43.9%, en esta elección para gobernador los resultados son claros y el próximo partido que gobernará Nuevo León es el PAN.

4.3.3 EFECTO DE DISEÑO DE LA ENCUESTA DE SALIDA

El efecto de diseño de la encuesta de salida se presenta en el cuadro siguiente:

la varianza del diseño muestral:

$$V(\hat{P}) = 0.0473\%, \quad (4.3.1)$$

la varianza de muestreo aleatorio simple:

$$V_{nias}(\hat{P}) = 0.0051\%, \quad (4.3.2)$$

dividiendo (4.3.1) entre (4.3.2) nos queda:

$$Deff = 9.20, \quad (4.3.3)$$

y el tamaño con una confianza de 95% y $\pm 5\%$ error absoluto:

$$n = n_{nias} Deff = 385 * 9.2 \cong 3542$$

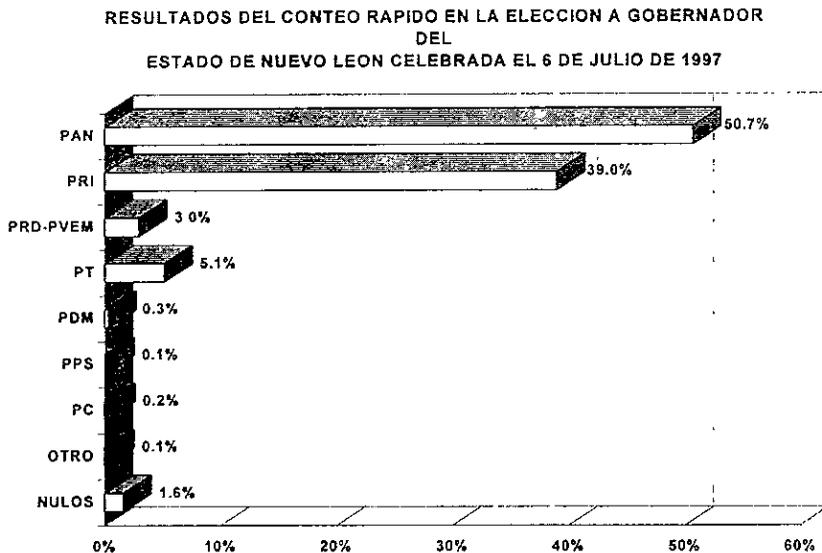
dividiendo n entre 30 (encuestas efectivas por sección) quedan 119 secciones electorales, es decir, para realizar una encuesta de salida con una confianza del 95% y un error absoluto de $\pm 5\%$, necesitamos realizar en 119 secciones electorales un total de $119 * 30 = 3570$ entrevistas.

4.4. CONTEO RAPIDO

4.4.1. RESULTADOS DEL CONTEO RAPIDO

Debido a que el conteo rápido es más preciso que la encuesta de salida y no existe el reactivo “no contestó”, se puede visualizar un resultado más cercano a la realidad y observar que el PAN obtuvo el triunfo en la jornada electoral de la elección para gobernador. Los resultados del conteo rápido se presentan en el siguiente gráfico 4.4.A:

Gráfico 4.4.A Resultados del conteo rápido en la elección para gobernador



Se puede confirmar con respecto a la encuesta de salida o postvoto que el ganador según el resultado del conteo rápido es el PAN con un 51.0%, el PRI obtiene un 39.0%, los resultados el conteo rápido se espera que sean un reflejo de lo que ocurrió al final de la jornada electoral, cabe destacar que aquí existe un reactivo que se llama “nulos”, estos votos son generalmente para aquellos que acudieron a votar y no votaron por ningún partido político, los que votaron por algún candidato no registrado, aquellos que votaron por dos o más partidos políticos o se salieron de la casilla a votar.

Los resultados del conteo rápido por estrato se presentan en el siguiente cuadro 4-13:

Cuadro 4-13: Resultados del conteo rápido de la elección del 6 de julio de 1997 para gobernador del Estado de Nuevo León según estrato

ESTRATO	PAN	PRI	PRD-PVEM	PT	PDM	PPS	OTRO	NULOS	TOTAL
MONTERREY	48.7%	40.0%	3.6%	5.6%	0.3%	0.1%	0.1%	1.6%	100.0%
GUADALUPE	53.5%	36.3%	3.3%	5.0%	0.2%	0.1%	0.1%	1.5%	100.0%
SAN NICOLAS	60.7%	29.0%	3.0%	5.5%	0.3%	0.1%	0.4%	1.0%	100.0%
APODACA Y G ESCOBEDO	43.8%	44.3%	2.7%	6.7%	0.3%	0.2%	0.3%	1.7%	100.0%
SAN PEDRO	67.3%	27.7%	1.5%	2.0%	0.8%	0.0%	0.1%	0.6%	100.0%
STA CATARINA	59.4%	30.1%	4.4%	4.1%	0.6%	0.0%	0.0%	1.4%	100.0%
OTROS	42.3%	50.0%	1.2%	3.8%	0.2%	0.2%	0.5%	1.8%	100.0%

Según estos resultados el PAN obtendría la victoria en Monterrey, Guadalupe, San Nicolás de los Garza, San Pedro Garza García y Santa Catarina, el PRI obtendría la victoria en el conjunto Apodaca y General Escobedo y el resto del Estado.

Los resultados presentados en el cuadro 4-13 no necesariamente son representativos por estrato, ya que el muestreo se diseñó para el conjunto estatal en la elección para gobernador y la cantidad de secciones electorales no podrían ser el adecuado, sin embargo, si nos puede dar una idea en algunas ocasiones sobre el comportamiento de éstos; en Monterrey se tomaron datos de 61 secciones electorales de la elección para gobernador, con esta cantidad de secciones electorales se está en condiciones de dar resultados confiables de Monterrey.

4.4.2 VARIANZA E INTERVALOS DE CONFIANZA

La varianza para el conteo rápido es de suma importancia para poder determinar la confiabilidad de un resultado, sobre todo si en la encuesta de salida se detectó un empate técnico, debido a que el conteo rápido tiene mayor confiabilidad que una encuesta de salida, es posible en algunas ocasiones poder dar un triunfador que no fue posible en la encuesta de salida; un ejemplo concreto en mi experiencia fue el municipio de Tonalá, Jalisco en la elección del 9 de Noviembre de 1997 a Presidente Municipal, el resultado de la encuesta de salida fue declarado empate técnico, pero al recibir los datos del conteo rápido se pudo decir quien sería el próximo Presidente Municipal; desde luego no es el caso del Estado de Nuevo León, donde la tendencia fue constante y la diferencia porcentual entre un partido político y otro fue amplia, además los intervalos de confianza no se intersectaron. Las varianzas del conteo rápido se presentan a continuación:

Cuadro 4-14: Varianza de los principales partidos políticos de la elección del 6 de julio de 1997 para gobernador de Nuevo León según estrato en el conteo rápido

ESTRATO	PAN	PRI	PRD-PVEM	PT
MONTERREY	0.0397%	0.0251%	0.0005%	0.0087%
GUADALUPE	0.0432%	0.0421%	0.0009%	0.0010%
SAN NICOLAS	0.0386%	0.0336%	0.0008%	0.0011%
ÁPODACA Y G. ESCOBEDO	0.0890%	0.0867%	0.0016%	0.0038%
SAN PEDRO	0.2792%	0.2060%	0.0011%	0.0035%
STA. CATARINA	0.1127%	0.1069%	0.0025%	0.0017%
OTROS	0.3315%	0.1730%	0.0070%	0.0517%
GLOBAL	0.0180%	0.0113%	0.0003%	0.0026%

Determinada la varianza se pueden estimar los intervalos de confianza por estrato y de manera global de los partidos políticos que obtuvieron la mayor cantidad de votos en la elección. Los intervalos de confianza se presentan a un 95%.

En el cuadro 4-15 se puede observar que el valor inferior del intervalo de confianza del PAN en el ámbito estatal es mayor con un 48.1% que el valor superior del intervalo de confianza del PRI con un 41.1%.

Cuadro 4-15: Límites de confianza por estrato y global de los principales partidos en la elección del 6 de julio de 1997 para gobernador de Nuevo León en el conteo rápido

ESTRATO	LIMITE	PAN	PRI	PRD-PVEM	PT	CV(VOTOS)
MONTERREY	INFERIOR	44.8%	36.9%	3.2%	3.8%	6.3%
	SUPERIOR	52.6%	43.1%	4.0%	7.4%	
GUADALUPE	INFERIOR	49.4%	32.3%	2.7%	4.4%	6.8%
	SUPERIOR	57.6%	40.3%	3.9%	5.6%	
SAN NICOLAS	INFERIOR	56.8%	25.4%	2.4%	4.9%	7.1%
	SUPERIOR	64.6%	32.6%	3.6%	6.1%	
APODACA Y G. ESCOBEDO	INFERIOR	38.0%	38.5%	1.9%	5.5%	19.1%
	SUPERIOR	49.6%	50.1%	3.5%	7.9%	
SAN PEDRO	INFERIOR	56.9%	18.8%	0.9%	0.8%	12.1%
	SUPERIOR	77.7%	36.6%	2.1%	3.2%	
STA CATARINA	INFERIOR	52.8%	23.7%	3.4%	3.3%	15.8%
	SUPERIOR	66.0%	36.5%	5.4%	4.9%	
OTROS	INFERIOR	31.0%	41.8%	0.0%	0.0%	18.5%
	SUPERIOR	53.6%	58.2%	2.8%	8.3%	
GLOBAL	INFERIOR	48.1%	36.9%	2.6%	4.1%	5.6%
	SUPERIOR	53.3%	41.1%	3.4%	6.1%	

Recordemos que las varianzas son válidas si el tamaño por estrato es mayor que 30 y el coeficiente de variación menor que 10% (pág 108). Cabe destacar que en el conteo rápido los coeficientes de variación son menores de 20%.

4.4.3 EFECTO DE DISEÑO DEL CONTEO RAPIDO

En el apartado 2.7 se presentó el efecto de diseño y las bondades de éste, en el presente apartado se expone el efecto de diseño y un tamaño de muestra bajo este modelo muestral para un conteo rápido. Sólo cabe recordar que el efecto de diseño se calcula para el partido político mayoritario, para este caso concreto fue el PAN.

La varianza del estimador del PAN, bajo el modelo muestral elaborado es:

$$V(\hat{p}_{PAN}) = 0.018\%, \quad (4.4.1)$$

la varianza bajo un modelo de muestreo aleatorio simple es:

$$V_{mas}(\hat{p}_{PAN}) = 0.000175\%, \quad (4.4.2)$$

dividiendo (4.4.1) entre (4.4.2) nos queda:

$$Deff(\hat{p}_{PAN}) = 103. \quad (4.4.3)$$

Para calcular el tamaño de muestra tomamos en cuenta lo siguiente: calculamos un tamaño para un muestreo aleatorio simple con una confianza del 95%, un error relativo de 10% y un $P = 47.1\%$, utilizando (2.1.12) nos queda:

$$n_0 = 432, \quad (4.4.4)$$

sustituyendo n_0 (4.4.4) en (2.1.11) para calcular n_{mas} nos queda:

$$n_{mas} = 432, \quad (4.4.5)$$

de esta manera multiplicamos (4.4.5) por (4.4.3) y nos queda:

$$n = n_{\text{max}} \text{Deff}(\hat{p}) = 44496,$$

pero si dividimos n entre el tamaño promedio de 649 electores por sección electoral en el Estado de Nuevo León, obtenemos la cantidad de secciones electorales para un igual diseño, así nos resulta:

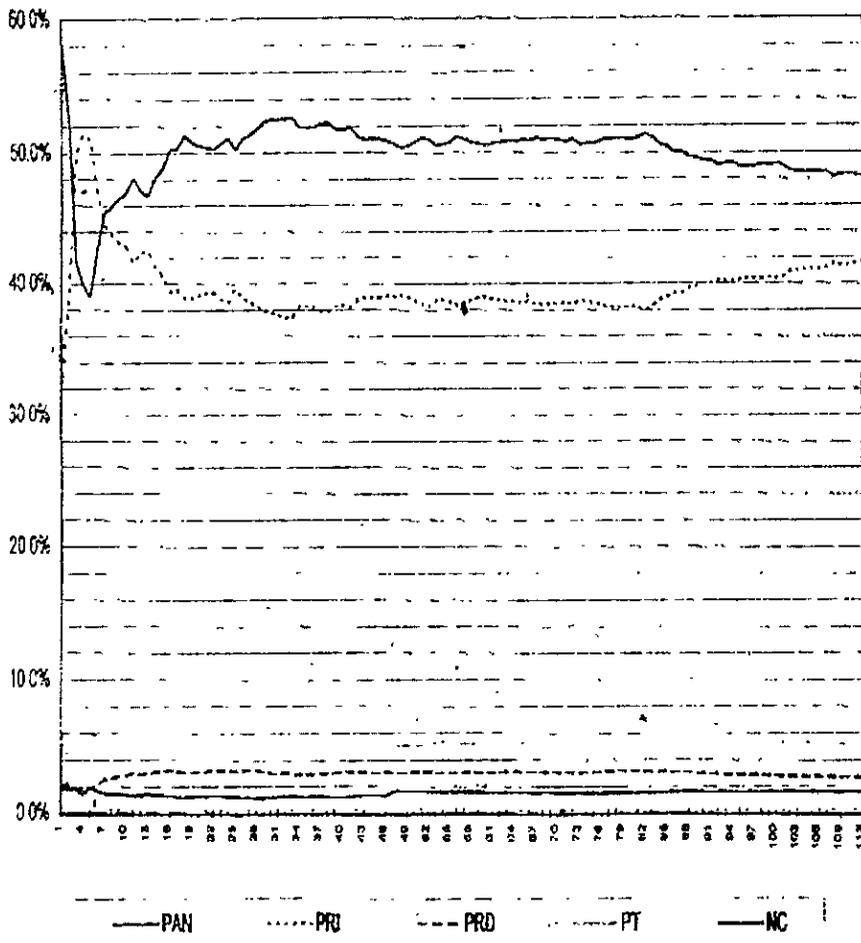
$$n = 70.$$

Setenta secciones electorales serían las que se necesitarían para el Estado de Nuevo León.

4.4.4 EVOLUCION DEL VOTO DEL CONTEO RAPIDO

Gráfico 4.4.B Evolución del voto según número de secciones electorales reportadas en la elección para gobernador del Estado de Nuevo León celebrada el 6 de julio de 1997

NUEVO LEON EVOLUCION EN LA DISTRIBUCION DEL VOTO (CONTEO RAPIDO) POR PARTIDO EN LAS ELECCIONES PARA GOBERNADOR



El gráfico 4.4.B determina la evolución de la intención del voto en el conteo rápido a través del número de secciones que se van registrando a través del tiempo, es decir, se obtiene un gráfico que mide el resultado acumulado con respecto a la cantidad de secciones electorales, cuando la curva comienza a tener una pendiente que tiende a cero y las curvas entre un partido político y otro no se intersectan, así se puede asegurar que este resultado difícilmente llegará a variar.

El gráfico 4.4.B indica la distribución porcentual de la intención del voto en el conteo rápido de los partidos políticos a través del número de secciones que se reportan y se registran, por ejemplo, al llegar la primera sección electoral se muestra que el PAN esta muy cerca de 60% y el PRI por debajo del 40%, la forma en que opera este gráfico acorde a la muestra es de la siguiente manera, es un estimador porcentual de una función de la cantidad de secciones registradas, por consiguiente se tiene:

$\hat{P}(n)$ = Porcentaje de electores que votaron por un partido de las n secciones reportadas,

de esta manera nos queda el primer estimador:

$$\hat{P}(1) = \frac{\text{\# de votos a favor de un partido determinado en el primer reporte}}{\text{\# de votos emitidos en el primer reporte}}, \quad (4.4.6)$$

Para construir el segundo estimador existen dos caminos: el primero de ellos es si ambas secciones electorales son del mismo estrato, y el segundo de es que sean de diferente estrato; recuérdese como se construyó la muestra en el conteo rápido, los estimadores dentro de cada estrato son autoponderados, pero entre los estratos es necesario utilizar el ponderador, entonces en el primer caso que son del mismo estrato queda como sigue:

$$\hat{P}(2) = \frac{\text{\# de votos a favor de un partido en el primero y segundo reportes}}{\text{\# de votos emitidos en el primero y segundo reportes}}, \quad (4.4.7)$$

pero si los reportes corresponden a diferentes estratos, se necesita generar un estimador por estrato equivalente a (4.4.6), y los dos ponderadores de los estratos reponderarlos, que no es otra cosa que hacer que sumen el 100% de manera proporcional entre ambos, así multiplicamos cada estimador (4.4.6) por su reponderador y de esta manera se genera (4.4.8):

Sea W_i, W_j los ponderadores de los estratos $i, j; i \neq j; i, j \leq 7$ entonces sean los reponderadores:

$$W'_i = \frac{W_i}{W_i + W_j}, \quad (4.4.8)$$

$$W'_j = \frac{W_j}{W_i + W_j}, \quad (4.4.9)$$

y

$$\hat{P}_i^1 = \text{estimador del primer reporte del estrato } i, \quad (4.4.10)$$

$$\hat{P}_j^2 = \text{estimador del segundo reporte del estrato } j, \quad (4.4.11)$$

de esta manera queda:

$$\hat{P}(2) = W'_i \hat{P}_i^1 + W'_j \hat{P}_j^2. \quad (4.4.12)$$

y así sucesivamente es como se construye la función del estimador con respecto al tiempo el gráfico, desde luego que con anterioridad se realizó un programa que de manera automática genera estos resultados puntuales y la curva presentada.

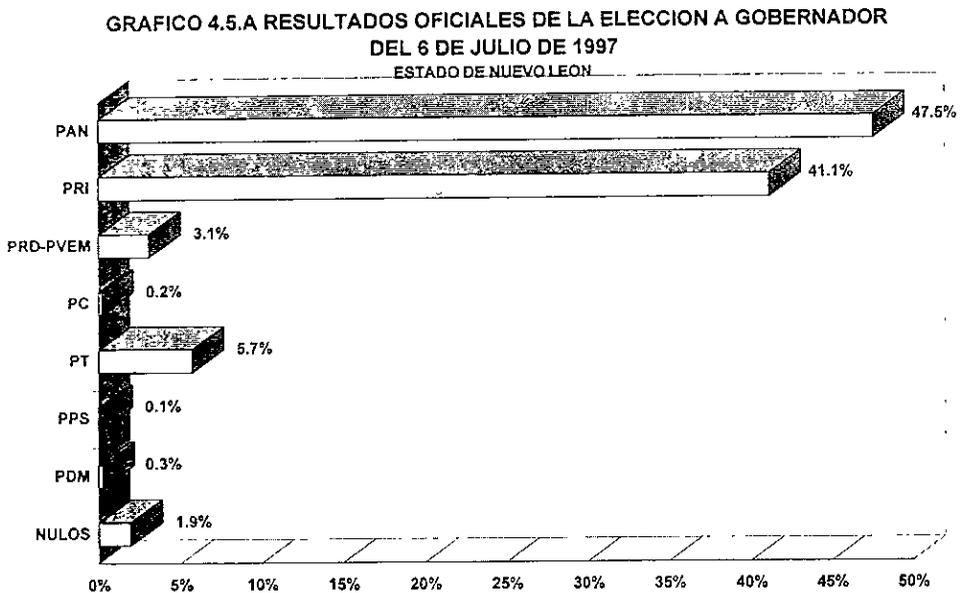
Después de la primera sección electoral reportada favorablemente hacia el PAN, éste comienza a bajar su porcentaje y el PRI inicia el ascenso del mismo por la llegada de secciones priístas y que no tienen problemas de comunicación, sobre todo telefónicos, la tendencia priísta se puede ver en los primeros siete reportes, con éstas secciones electorales no se está en

condiciones para poder ver una curva de pendiente cero, al menos en este caso de Nuevo León. Del séptimo al décimo reporte se llega al empate, por lo que se puede observar que comienzan a llegar secciones con preferencia hacia el PAN y de manera visual a partir de la décima sección supera el PAN de manera permanente al PRI, esto se debe a que las secciones que comienzan a llegar son urbanas y favorecen al PAN; las secciones urbanas en conjunto generalmente son reportadas primero que las rurales, debido a los medios de comunicación; por eso entre los reportes 10 y 28 siguen llegando secciones urbanas que favorecen al PAN, de allí que en este intervalo la diferencia sigue creciendo hasta pasar al intervalo del reporte 28 al reporte 82 donde se observa que las pendientes de las curvas son muy cercanas a cero, lo que indica que la diferencia entre el partido de mayoría, el PAN, y su cercano competidor, el PRI, se mantiene; en el último intervalo que es de la sección 82 hasta el resto de los reportes se observa que la tendencia hace que la diferencia entre un partido y otro se cierre, las últimas secciones en reportar fueron rurales de zonas alejadas y sin comunicación telefónica, las secciones con estas características favorecen al PRI, pero en como cada ocasión que aumenta el número de reportes ocasiona que un reporte nuevo sea menos significativo hacia la función del estimador y la cantidad de secciones eran considerables, el resultado del triunfo del PAN era irreversible.

4.5. RESULTADOS OFICIALES

4.5.1. PRESENTACION DE LOS RESULTADOS OFICIALES

Los resultados oficiales que se presentan en el gráfico 4.5.A son los publicados por la Comisión Estatal Electoral del Estado de Nuevo León en la elección para gobernador del estado del día 6 de julio de 1997.



El Acta de resultados oficiales fue publicada el día 1 de Agosto de 1997 en el Periódico Oficial del Gobierno Constitucional del Estado Libre y Soberano de Nuevo León de acuerdo al artículo 216 de la Ley Electoral del Estado. Se publica el Acta de la sesión de cómputo y declaratoria de validez celebrada por la Comisión Estatal Electoral de los días 11 al 15 de julio de 1997, firman el Acta Ing. Ramón de la Peña Manrique y Lic. Gilberto Rogelio Villarreal de la Garza, Presidente y Secretario de la Comisión Estatal Electoral.

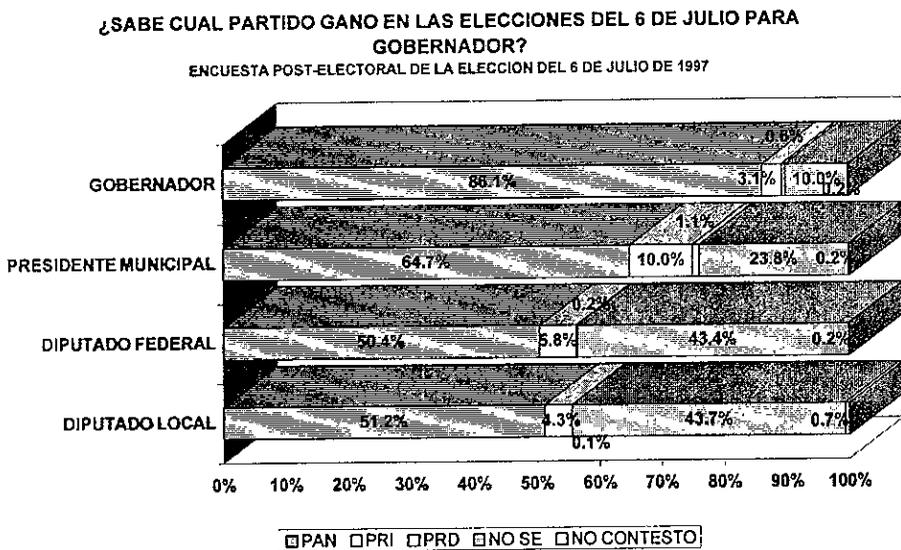
4.6. ENCUESTA POST-ELECTORAL

4.6.1 RESULTADOS DE LA ENCUESTA POST-ELECTORAL

A continuación se presentan algunos gráficos del estudio post-electoral, el objetivo de esta sección es la presentación de resultados, sólo se presentan algunos cruces y variables demográficas, los resultados de la intención de voto se comparan en el capítulo 5 con respecto a las demás encuestas, el conteo rápido y los resultados oficiales. El levantamiento de la encuesta se hizo los días 12 y 13 de julio, una semana después del día de la elección; el total de casos fue de 1,209 casos distribuidos en 80 secciones electorales.

De esta manera se visualizarán los gráficos uno a uno.

Gráfico 4.6.A Conocimiento sobre el triunfador en la jornada electoral en Nuevo León



Se puede ver que la mayoría de la población tiene un mayor conocimiento de la elección para gobernador que los demás puestos de elección popular, el 86.1% contestó bien la pregunta de quien fue el partido ganador en la elección de gobernador en el Estado de Nuevo León.

Según la encuesta post-electoral donde se preguntó si habían acudido a votar, mencionaron un 72.9% que sí y el resto que no, comparada con el 63.8% de votación oficial tiene una diferencia de 9.1%. Sobre la pregunta de por cual partido político votó el pasado 6 de julio de 1997 el 57.8% mencionó haber votado por el PAN y el 33.6% por el PRI, probablemente la mayoría del electorado al saber que el PAN había obtenido el triunfo no quiso mencionar que había votado por otro partido, sin embargo, lo que aquí se mide es la percepción de la gente en el momento del levantamiento de la encuesta, se puede ver en el gráfico 4.6.D que aquellas personas que respondieron que votaron por el PRI, es por que siempre votan por el mismo, el llamado voto duro y en el gráfico 4.6.C aquellas personas que respondieron que votaron por el PAN lo hicieron buscando un cambio.

Gráfico 4.6.B Partido por el que votaron en la elección de Nuevo León

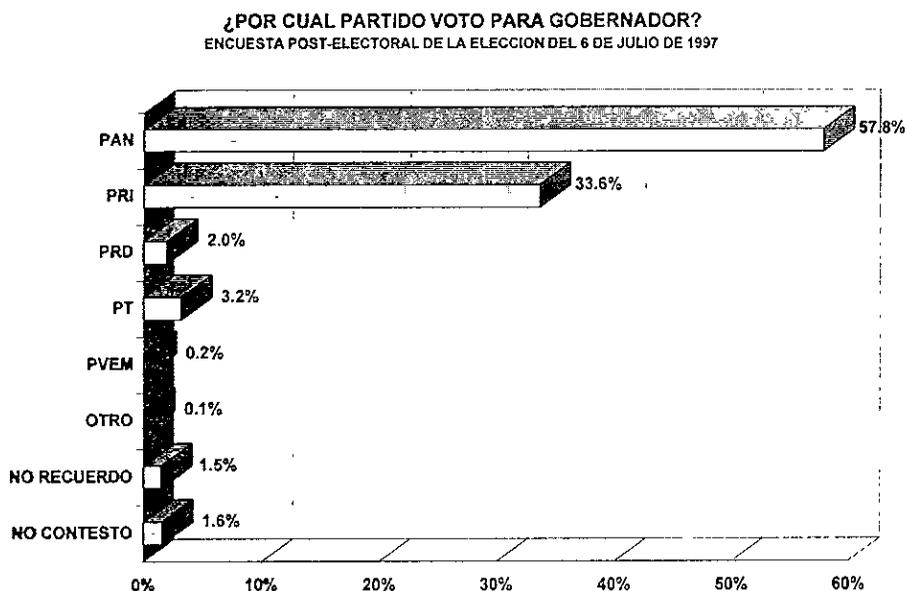
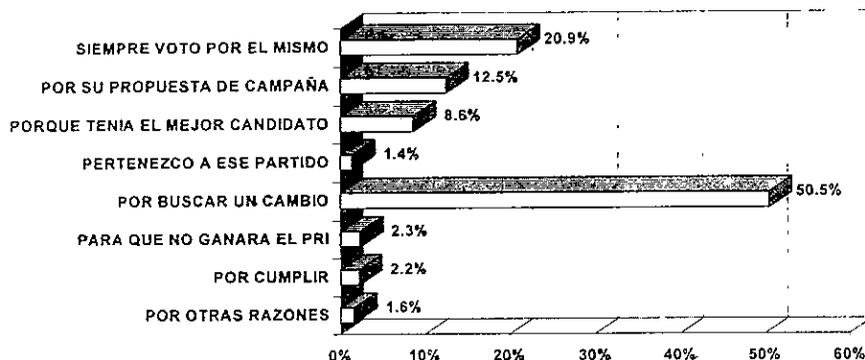


Gráfico 4.6.C Razón por la cual votaron por el PAN en Nuevo León (699 casos)

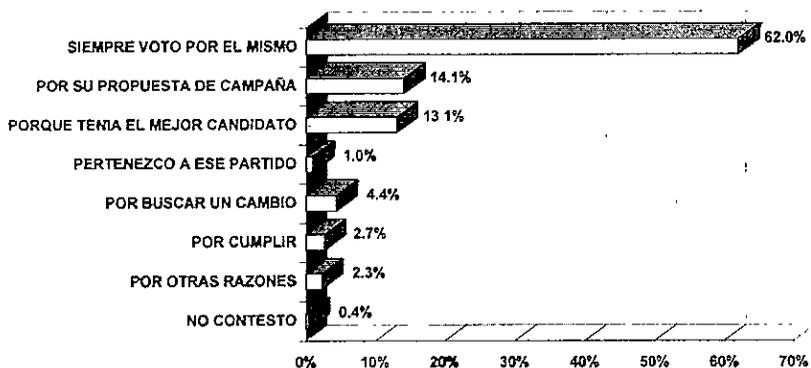
¿POR QUE RAZON VOTO POR ESE PARTIDO PARA GOBERNADOR?
ENCUESTA POST-ELECTORAL DE LA ELECCION DEL 6 DE JULIO DE 1997
(57.8% DE LOS CIUDADANOS QUE VOTARON POR EL PAN)



La diferencia del voto por el PAN y el PRI es que votaron por un cambio los del PAN y el voto característico del PRI es el voto duro.

Gráfico 4.6.D Razón por la cual votaron por el PRI en Nuevo León (406 casos)

¿POR QUE RAZON VOTO POR ESE PARTIDO PARA GOBERNADOR?
ENCUESTA POST-ELECTORAL DE LA ELECCION DEL 6 DE JULIO DE 1997
(33.6% DE LOS CIUDADANOS QUE VOTARON POR EL PRI)

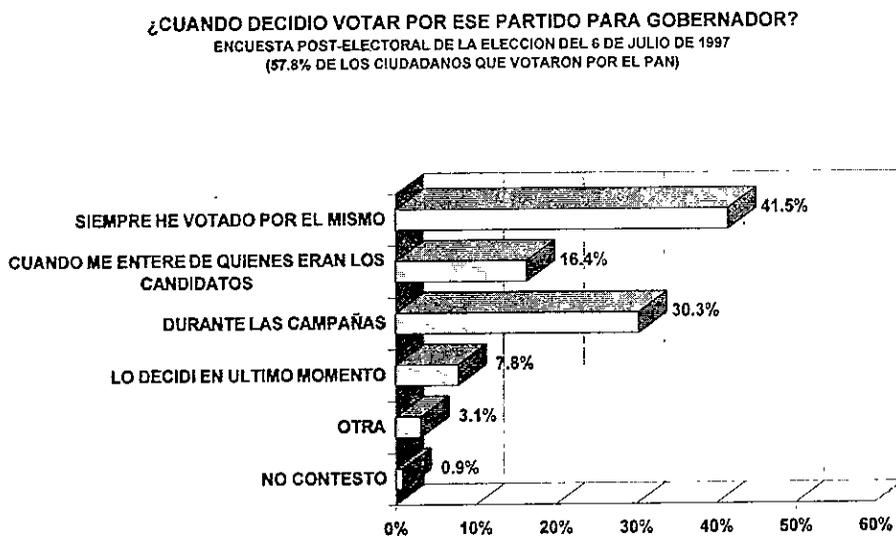


Si bien conocemos las causas por las que el electorado se decidió votar por cierto partido político, no conocemos aún en que tiempo lo hizo, tal vez el tiempo en el que un elector se decide es el mejor momento para influir en su decisión, esta es una herramienta para saber el impacto que tuvo cada una de las estrategias de campaña y se muestra que fue lo que impactó más a los electores.

Una vez más se continua con la presentación de los gráficos que nos explican en que momento el elector se decidió votar por ese partido.

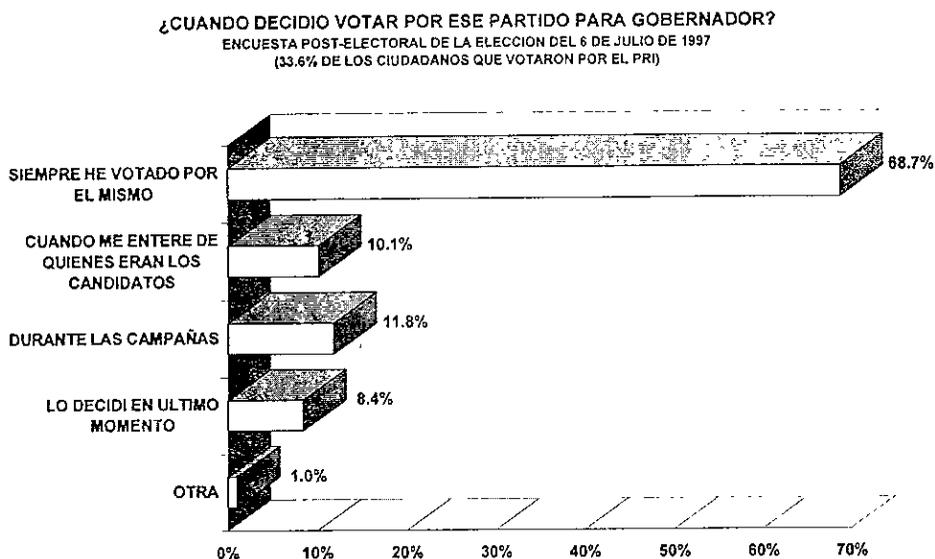
De esta manera se pueden ver los gráficos 4.6.E y 4.6.F.

Gráfico 4.6.E Tiempo en el que decidieron votar por el PAN en Nuevo León (699 casos)



Los que votaron por el PAN mencionan que siempre votan por el mismo con un 41.5%, sin embargo, existe un 30.3% importante que dicen que durante las campañas establecieron su preferencia, los que lo hicieron por el PRI mencionan que siempre votan por el mismo, una vez más se hace mención del llamado voto duro.

Gráfico 4.6.F Tiempo en el que decidieron votar por el PRI en Nuevo León (406 casos)



Para conocer las causas de los que no fueron a votar, ¿por qué no lo hicieron?, se generó el gráfico 4.6.G

Las tres principales causas por la que algunos ciudadanos no acudieron a votar son: con un 33.9% no tenían credencial, el 21.5% afirmó que no tenía tiempo y un 14.2% mencionó que no acudió a votar por que se encontraba de viaje.

Si se desea saber por quien hubieran votado las personas que no lo hicieron, se generó el gráfico 4.6.H, en que se les preguntó que ¿de haber ido a votar por quien lo hubieran hecho?, donde se establece que la un 40.8% hubiera votado por el PAN, seguido por el PRI con un 34.5%.

Gráfico 4.6.G Causas por las cuales la población no acudió a votar en el Estado de Nuevo León en la elección para gobernador (279 casos)

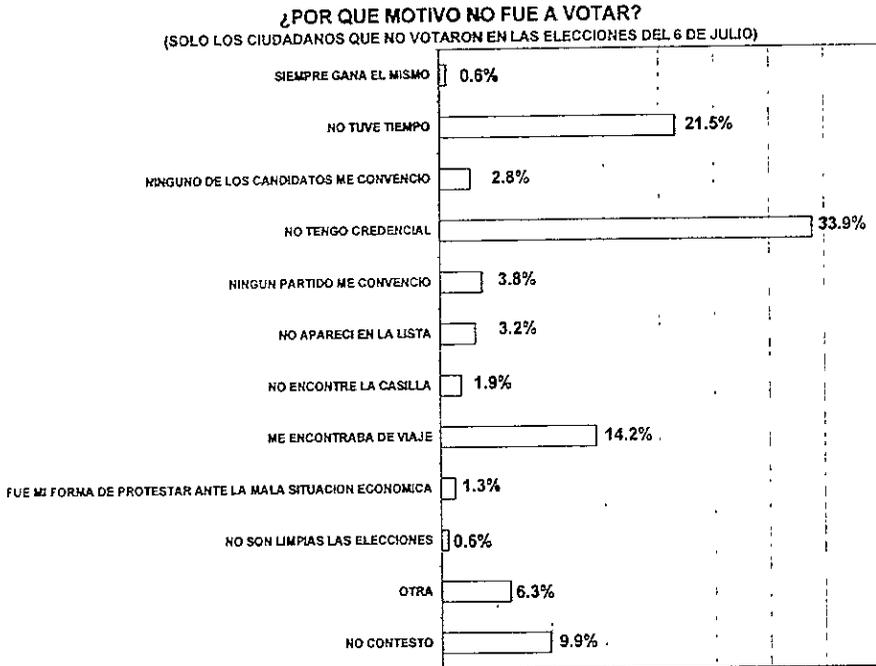
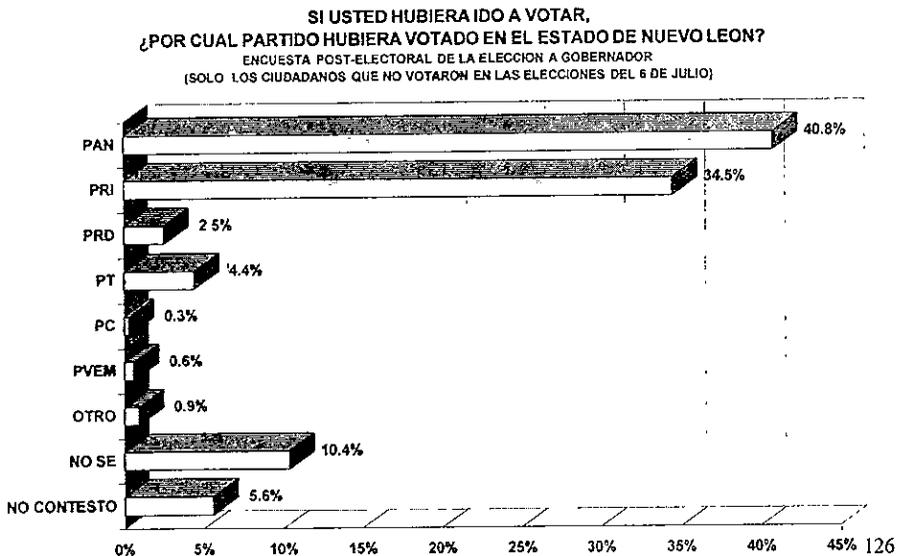


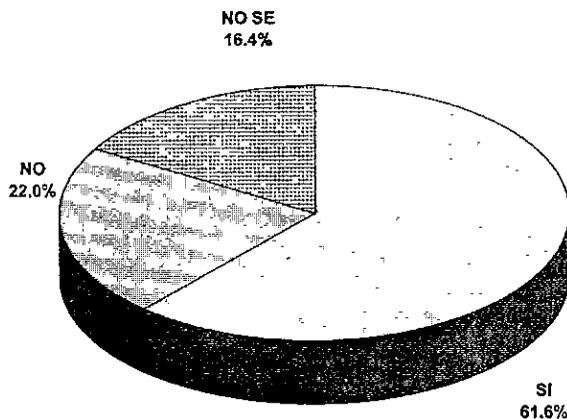
Gráfico 4.6.H Intención de voto de los que no acudieron a votar (279 casos)



En el gráfico 4.6.I se puede ver que la mayoría de la ciudadanía cree que las elecciones para gobernador fueron limpias y transparentes.

Gráfico 4.6.I Creencia sobre la transparencia de las elecciones en Nuevo León

¿CREE QUE LAS PASADAS ELECCIONES FUERON LIMPIAS Y TRANSPARENTES?
ENCUESTA POST-ELECTORAL DE LA ELECCION DEL 6 DE JULIO DE 1997 A GOBERNADOR



Según el gráfico 4.6.J, el porcentaje de ciudadanos que mencionaron pertenecer a algún partido político es del 6.5%, para saber a que partidos políticos pertenecen está el gráfico 4.6.K; para medir el impacto de las campañas políticas que se realizaron sobre los electores y si estas influyeron en su decisión final pueden visualizarse los gráficos 4.6.L, 4.6.M y 4.6.N.

Para conocer el nivel de aceptación de la población hacia el Presidente de México sobre el apoyo y su actitud en el proceso de la elección para gobernador el día 6 de julio de 1997 en Nuevo León se generó el gráfico 4.6.O, donde recibió una calificación promedio de 7.4.

Gráfico 4.6.J Porcentaje de ciudadanos dentro de algún partido político en Nuevo León

¿PERTENECE A ALGUN PARTIDO POLITICO?
ENCUESTA POST-ELECTORAL DE LA ELECCION DEL 6 DE JULIO DE 1997
A GOBERNADOR

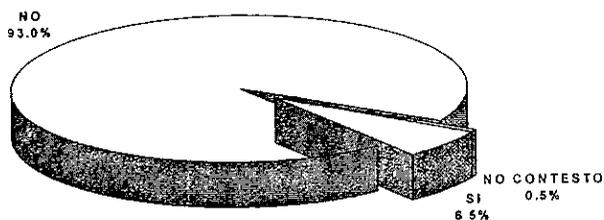


Gráfico 4.6.K Cruce de los que pertenecen a algún partido político en el Estado de Nuevo León, a cual partido político pertenecen? (79 casos)

¿A CUAL PARTIDO POLITICO PERTENECEN?
ENCUESTA POST-ELECTORAL DE LA ELECCION DEL 6 DE JULIO A GOBERNADOR EN EL
ESTADO DE NUEVO LEON
(6.5% DE LOS CIUDADANOS QUE PERTENECEN ALGUN PARTIDO POLITICO)

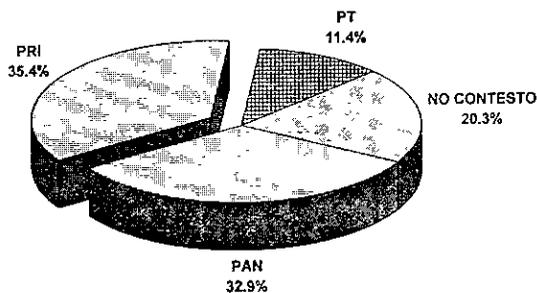


Gráfico 4.6.L Opinión sobre las campañas políticas en Nuevo León

¿QUE OPINA DE LAS CAMPAÑAS QUE DESARROLLARON LOS PARTIDOS POLITICOS PARA LAS ELECCIONES DEL 6 DE JULIO?

ENCUESTA POST-ELECTORAL DE LA ELECCION A GOBERNADOR

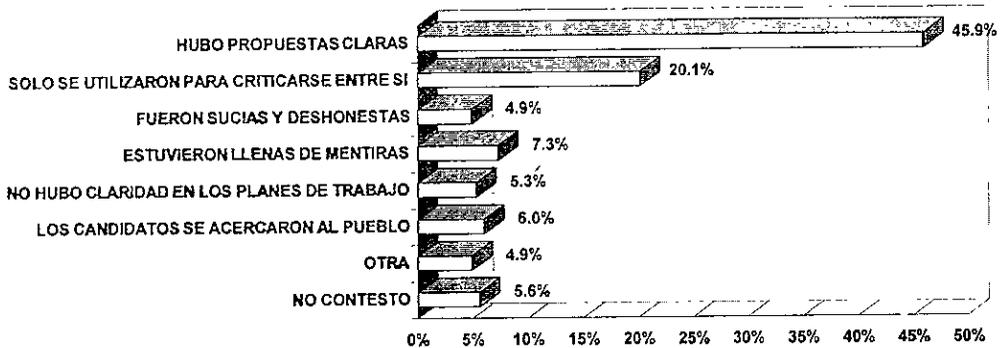


Gráfico 4.6.M Influencia de las campañas políticas en los electores de Nuevo León

EN SU DECISION DE VOTO PARA EL PASADO 6 DE JULIO, ¿INFLUYERON LAS CAMPAÑAS QUE LLEVARON A CABO LOS PARTIDOS?

ENCUESTA POST-ELECTORAL DE LA ELECCION A GOBERNADOR

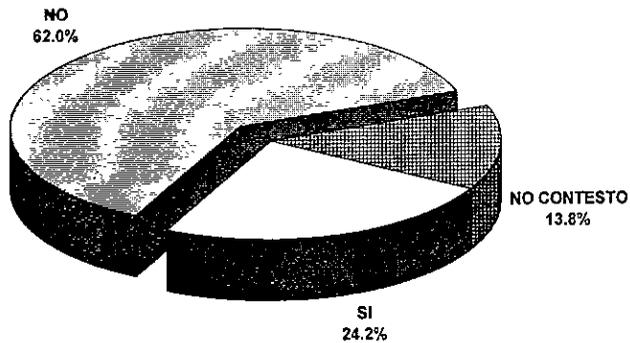


Gráfico 4 6.N Cruce de cómo influyeron las campañas en los que influyó (293 casos)

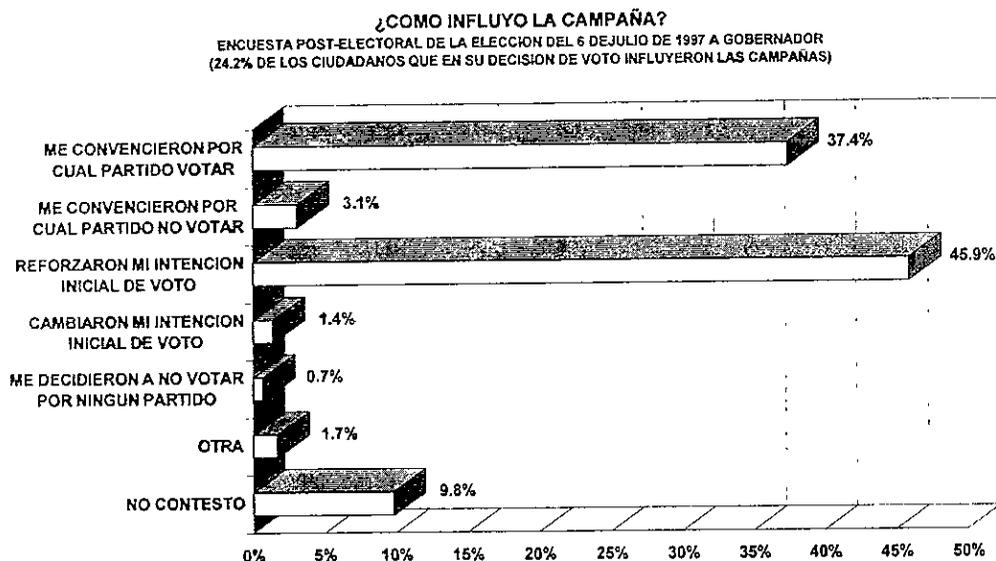
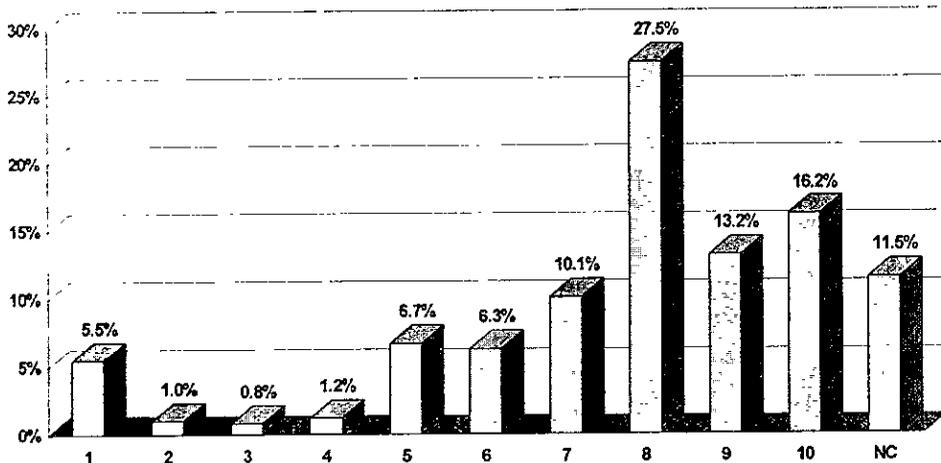


Gráfico 4.6.O Calificación otorgada al Presidente por el apoyo a la elección

¿QUE CALIFICACION LE PONDRIA A LA ACTTUD DE APOYO A LA DEMOCRACIA DEL PRESIDENTE DE LA REPUBLICA EN EL PROCESO ELECTORAL DEL 6 DE JULIO?
ENCUESTA POST-ELECTORAL DE LA ELECCION A GOBERNADOR EN EL ESTADO DE NUEVO LEON
MEDIA ARITMETICA 7.4



Para conocer las propuestas de la población en general hacia los candidatos, se les preguntó, que de tener la oportunidad de platicar con los triunfadores de las elecciones pasadas, ¿qué les sugeriría para que mejoraran las cosas en su comunidad?, las opciones manifestadas fueron que: exista más trabajo con un 21.6%, más honestidad con un 19.7% y más comunicación con el pueblo un 11.4%, sin embargo, el reactivo que más se contestó y que la población desea fue que cumpla lo que promete con un 33.2%.

Otra de las preguntas que se le realizó a la población fue sobre las medidas económicas que el gobierno aplicó, ¿a que sector de la población creen que debe beneficiar?, a lo que respondieron mayoritariamente que a los trabajadores; en contraste, la población opina que las decisiones económicas favorecieron a los empresarios y que el principal problema que las próximas autoridades deben resolver debe ser la seguridad pública.

Por último en esta encuesta post-electoral se valora si las encuestas influyeron en la decisión de los ciudadanos en el Estado de Nuevo León, su nivel de aceptación dentro de la ciudadanía y las causas por las cuales la ciudadanía las aprueba o desaprueba según sea el caso; asimismo, conocer el porcentaje de la población que se entera de resultados de las encuestas y conocer si éstas influyeron en su decisión final sobre su preferencia electoral.

Estos argumentos se pueden visualizar en las gráficas 4.6.P a la 4.6.T, donde se comprueba que las encuestas no son leídas por la mayoría de la población, sólo el 20.2% se enteró de los resultados de una encuesta, además, según el cruce de los que se enteraron de alguna encuesta (el 20.2%) sólo influyó en el 13.5% para decidir su preferencia electoral, de esta manera solamente influyó en el 2.72% de la población; cabe destacar que la cantidad de encuestas para estas conclusiones son 207, sin embargo, las conclusiones sobre las encuestas aquí mostradas no tienen la intención de representatividad, pero si darnos una idea importante sobre el tema.

Gráfico 4.6.P Acuerdo de la ciudadanía en las encuestas electorales en Nuevo León

¿ESTA DE ACUERDO EN QUE SE LLEVEN A CABO ENCUESTAS ELECTORALES?
ENCUESTA POST-ELECTORAL DE LA ELECCION DEL 6 DE JULIO DE 1997 A GOBERNADOR

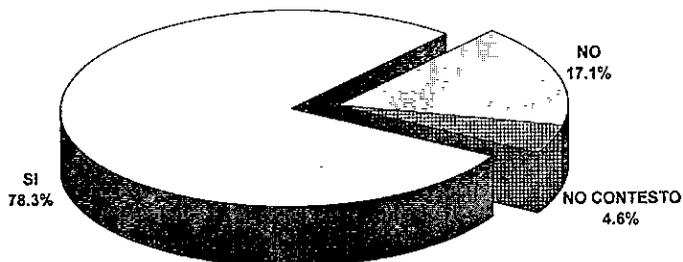


Gráfico 4.6 Q Cruce de los que están de acuerdo en encuestas electorales (947 casos)

¿POR QUE ESTAN DE ACUERDO EN ENCUESTAS ELECTORALES?
ENCUESTA POST-ELECTORAL DE LA ELECCION DEL 6 DE JULIO DE 1997 A GOBERNADOR
(78.3% DE LOS CIUDADANOS QUE ESTAN DE ACUERDO EN QUE SE LLEVEN A CABO ENCUESTAS)

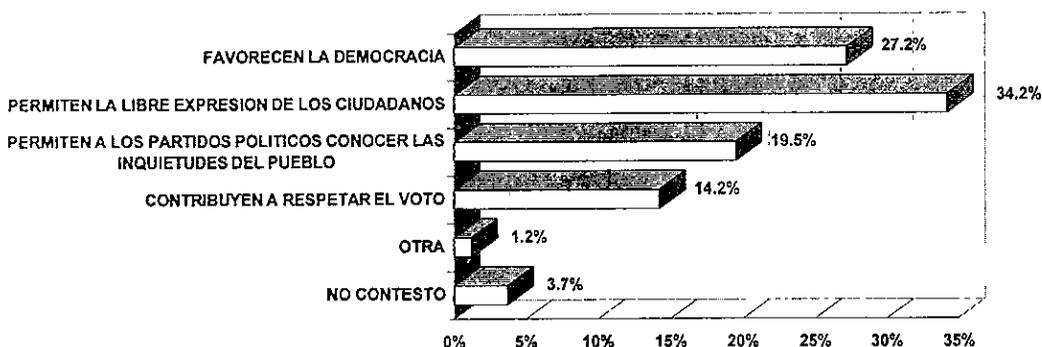


Gráfico 4.6.R Cruce de los que no están de acuerdo en encuestas electorales en el Estado de Nuevo León (207 casos)

¿POR QUE NO ESTA DE ACUERDO EN ENCUESTAS ELECTORALES?
ENCUESTA POST-ELECTORAL DE LA ELECCION DEL 6 DE JULIO DE 1997 A GOBERNADOR
(17.1% DE LOS CIUDADANOS QUE NO ESTAN DE ACUERDO EN QUE SE LLEVEN A CABO ENCUESTAS)

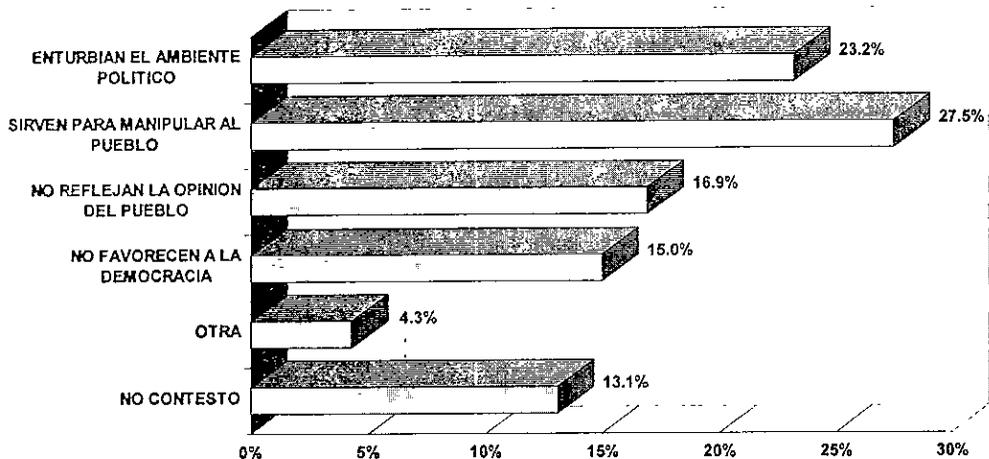


Gráfico 4.6 S Nivel de conocimiento de la ciudadanía hacia resultados de encuestas en el Estado de Nuevo León

¿SE ENTERO DE LOS RESULTADOS DE ALGUNA ENCUESTA?
ENCUESTA POST-ELECTORAL DE LA ELECCION DEL 6 DE JULIO DE 1997 A GOBERNADOR

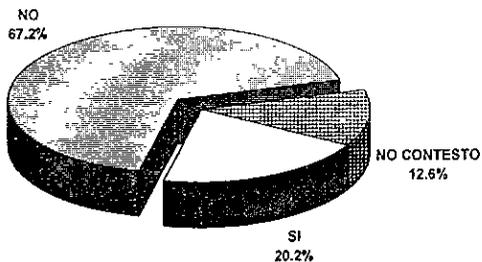
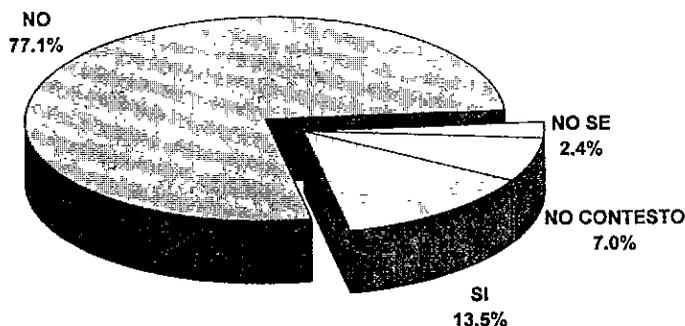


Gráfico 4.6.T Cruce Influencia sobre la decisión en los que vieron encuestas electorales en el Estado de Nuevo León (244 casos)

¿LOS RESULTADOS DE LAS ENCUESTAS INFLUYERON EN SU DECISION DE VOTO EL 6 DE JULIO DE 1997?

ENCUESTA POST-ELECTORAL DE LA ELECCION A GOBERNADOR
(20.2% DE LOS CIUDADANOS QUE SE ENTERARON DE LOS RESULTADOS DE ALGUNA ENCUESTA)



4.6.2 COMPARATIVOS DE LA ENCUESTA POST-ELECTORAL

Las variables que se analizaron en el apartado 4.2.3 donde se veían las diferencias entre reales y estimadas; dos de ellas cuotas y la variable “nivel de instrucción” en la que se podrían detectar posibles sesgos del muestreo. Estas variables se presentan ahora para la encuesta post-electoral.

Cuadro 4-16. Porcentajes de género mayor de 18 años según Conteo95 publicado por el Inegi y encuesta post-electoral de la elección del 6 de julio de 1997 para gobernador de Nuevo León

	HOMBRES	MUJERES
ENCUESTA	49.50%	50.50%
REAL	49.20%	50.80%
DIFERENCIA	0.26%	-0.26%

La diferencia de porcentajes según género en la encuesta post-electoral con la real no es significativa, el operativo de campo cumplió con esta cuota, ya que los porcentajes asignados al operativo fueron respetados.

Cuadro 4-17. Porcentajes de grupos de edad mayor de 18 años según Censo 95 publicado por el Inegi y encuesta post-electoral de la elección del 6 de julio de 1997 para gobernador de Nuevo León

EDAD	POB > 18	ENCUESTA	DIF
18-24	27.37%	26.60%	0.77%
25-34	27.51%	27.60%	-0.09%
35-44	18.13%	18.29%	-0.16%
45 y >	26.98%	27.52%	-0.54%

La otra cuota aplicada en la muestra del Estado de Nuevo León fue grupos de edad, las diferencias son pequeñas y también no significativas, cada cuota tiene menos de un punto porcentual de diferencia (ver cuadro 4-17), en este reactivo los encuestadores también cumplieron con lo señalado.

En el cuadro 4-18, a diferencia de las otras dos variables que se les indicó a los encuestadores con instrucciones precisas de obtener tal cantidad de encuestas con criterios de género y edad, y se visualizaron en los cuadros 4-16 y 4-17 las diferencias, en el cuadro 4-18 se presenta una variable que a los encuestadores no se les indicó la tomaran en cuenta como cuota. Se observa que la muestra no tiene sesgo considerable para el nivel de instrucción, la opción con mayor diferencia es para aquellos que tienen un nivel de instrucción de primaria, en el cuadro se observa una subestimación de 4.14%.

Cuadro 4-18. Porcentajes de grupos según nivel de instrucción de mayores de 18 años según Censo 95 publicado por el Inegi y encuesta post-electoral de la elección del 6 de julio de 1997 para gobernador de Nuevo León.

NIVEL DE INSTRUCCIÓN	POST-ELECTORAL	OFICIAL	DIF
SIN ESTUDIOS	4.57%	6.50%	-1.93%
PRIMARIA	31.75%	35.89%	-4.14%
SECUNDARIA	25.77%	23.23%	2.54%
BACHILLERATO Y TEC.	23.31%	21.06%	2.25%
LICENCIATURA	14.27%	11.78%	2.49%
POSTGRADO	0.34%	1.54%	-1.20%

4.6.3 VARIANZA E INTERVALO DE CONFIANZA DE POR CUAL PARTIDO POLITICO VOTO

En este punto del trabajo, se presenta la varianza de ¿por cual partido político votó en las elecciones del 6 de julio?, las varianzas que se presentan son por estrato y en el ámbito estatal.

Cuadro 4-19: Varianzas de la encuesta post-electoral por estrato de la elección del 6 de julio de 1997 para gobernador de Nuevo León.

ESTRATO	PAN	PRI	PRD	PT
1	2.3577%	0.0326%	0.0259%	0.0150%
2	0.6390%	0.2033%	0.0049%	0.0693%
3	0.0596%	0.4607%	0.0246%	0.0263%
4	0.2816%	1.0573%	0.1046%	0.0000%
5	0.4214%	0.0669%	0.0186%	0.0180%
6	0.2076%	0.1597%	0.0000%	0.0306%
7	0.4169%	0.0345%	0.0145%	0.0173%
8	0.2826%	0.2585%	0.0052%	0.0715%
9	0.3181%	0.3917%	0.0149%	0.0158%
10	1.5917%	0.2490%	0.0213%	0.0162%
11	0.3073%	0.7389%	0.0000%	0.0000%
12	0.6622%	0.0772%	0.0098%	0.0098%
TOTAL	0.0690%	0.0198%	0.0012%	0.0025%

El estrato que mayor varianza tiene hacia el PAN es el 1, en general los estratos en que se utilizó un trietápico son los que mayor varianza tienen, a excepción del estrato 12 que comprende parte del municipio de Guadalupe donde se utilizó un bietápico y tiene alta variabilidad, el estrato que menor variabilidad tiene hacia el PAN es el 3, que es parte del municipio de San Nicolás de los Garza. El estrato que mayor variabilidad tiene hacia el PRI, es el estrato 4, que es el municipio de General Escobedo, y los estratos que menor variabilidad tienen hacia el PRI son el 7 que comprende parte del municipio de Monterrey y el 1 que comprende los municipios de Lampazos de Naranjo, San Pedro Garza García y Santa Catarina.

Los intervalos de confianza a un 95% para el PRI y el PAN se presentan en el cuadro siguiente:

Cuadro 4-20 Intervalos de confianza de por cual partido político votó en el ámbito estatal en la encuesta post-electoral de la elección para gobernador en Nuevo León celebrada el 6 de julio de 1997.

PARTIDO	PARAMETROS		INTERVALO DE CONFIANZA (95%)	
	MEDIA	VARIANZA	INFERIOR	SUPERIOR
PAN	57.8%	0.0690%	52.7%	62.9%
PRI	33.6%	0.0198%	30.8%	36.4%

En general se puede observar que estrato a estrato además del ámbito estatal el PAN mantuvo una alta variabilidad, esta variabilidad puede verse reflejada en los intervalos de confianza.

4.6.4 EFECTO DE DISEÑO DE LA ENCUESTA POST-ELECTORAL

Para la construcción del efecto de diseño se utilizó la varianza del partido político mayoritario. Las varianzas con los diferentes modelos se presentan a continuación:

La varianza del diseño muestral elaborado:

$$V(\hat{p}) = 0.0690\%, \quad (4.6.1)$$

la varianza del muestreo aleatorio simple:

$$V_{n_{max}}(\hat{p}) = 0.0204\%, \quad (4.6.2)$$

dividiendo (4.6.1) entre (4.6.2) nos queda:

$$Deff = 3.39, \quad (4.6.3)$$

utilizando (4.6.3) calculamos el tamaño de muestra con una confianza de 95% y un error absoluto de $\pm 5\%$.

$$n = n_{max} Deff = (375) * (3.39) \cong 1272. \quad (4.6.4)$$

5.- CONCLUSIONES Y COMENTARIOS

5.1 DE LOS PROCESOS ELECTORALES

En la introducción se menciona la importancia de las encuestas electorales en la época actual de México. Se mencionó también que éstas gozan en la mayoría de los casos de buena precisión, lo cual se puede visualizar en los resultados de la sección 5.2, en que la encuesta preelectoral, encuesta de salida o postvoto, conteo rápido y la encuesta post-electoral tienen un acercamiento aceptable con el resultado oficial. Sobre las encuestas electorales, una encuesta preelectoral es aquella que se realiza previo a una elección oficial que nos mide los parámetros de intención de voto, animadversión (rechazo hacia el partido) y la simpatía por los candidatos principalmente; la encuesta de salida o postvoto es aquella que se realiza a las personas después de emitir su voto en la casilla electoral respectiva, se le llama encuesta de salida porque se encuesta a las personas a la salida de la casilla, o postvoto porque se entrevista a las personas después de haber emitido su voto; el conteo rápido es aquel que mediante la recopilación directa de la publicación de resultados en la casilla electoral después de haber terminado la jornada electoral se genera un resultado muy parecido al oficial, como no existen “mentirosos” y también se elimina el reactivo no contestó, éste tiende a ser más preciso que la encuesta de salida; la encuesta post-electoral es la que se realiza después del día de la elección donde los tópicos importantes a conocer son porque los electores votaron por determinado partido.

Según la encuesta preelectoral en el estado de Nuevo León la gente estaba dividida en dos grandes bloques, el voto duro que generalmente favorece al PRI y los que querían un cambio, en la realidad se puede obtener como conclusión que la ciudadanía hasta el momento del levantamiento de la encuesta preelectoral votaría por la oposición por un cambio y no por las propuestas de trabajo de la misma, además, se pudo confirmar que la gente en la actualidad tiene mayor sensibilidad hacia el candidato que hacia el partido, este tipo de tendencias ha estado sucediendo en México en los últimos tiempos, pero al analizar la encuesta preelectoral nos damos cuenta que la imagen de ambos candidatos era muy similar, de allí la sencilla conclusión de que la gente votó por un cambio. En una encuesta de este tipo estuvimos en condiciones de generar cruces, que es cuando a una pregunta tiene cierta característica de otra, ésta puede ser sexo, edad u otra pregunta. De cualquier manera, la encuesta preelectoral daba como virtual triunfador al PAN, el cruce

de los indecisos decían que si las personas indecisas no votasen por el PAN, tampoco éstas lo harían por el PRI, de hecho era probable que estas no votarían por el PRI lo que hacía suponer que la distancia entre el PAN y el PRI tendería a ampliarse.

Después de haber hecho un análisis de la encuesta preelectoral, en la que el triunfo aparente sería competido pero hacia el PAN, se realiza en pleno día de la elección la encuesta de salida o postvoto, allí se define por quien votó el electorado que acudió a votar, se pudo constatar que en la elección la distancia entre el PRI y el PAN fue más amplia que la encuesta preelectoral, sin embargo, existe una alta respuesta de los que no contestaron a la pregunta por quien votó en la encuesta de salida, el argumento más fuerte generalmente es que el voto es secreto, desde luego que a la persona entrevistada se le debe garantizar el anonimato y ese punto es respetable, ahora la pregunta es hacia donde se distribuirían las personas que no contestaron, existen dos hipótesis sencillas, la primera es hacer uso de la encuesta preelectoral y saber por quien no votarían los que no contestan, pero se enfrentaría a la pregunta de si el universo de los que no contestan en la encuesta preelectoral y los que no contestaron en la encuesta de salida o postvoto son los mismos, la otra opción más sencilla y que ha arrojado buenos resultados es distribuir de manera proporcional los que no contestaron por que partido político lo hicieron, sobre los que sí lo hicieron, éste es un modelo sencillo. También es raro ver una encuesta en la que en el primer corte varía considerablemente con el segundo corte y éste altere los resultados, en esta tesis se puede ver que los resultados en cada uno de los cortes la intención del voto se mantuvo estable, después de terminada la jornada electoral, llega el momento de tomar decisiones fuertes, en casos en que las elecciones son competidas es necesario tomarlas de manera responsable y en el corto plazo, pero en el caso de la encuesta de salida del estado de Nuevo León no hubo necesidad de ello, ya que la elección tuvo un claro comportamiento con una amplitud suficiente para poder decir los resultados.

Llegada la hora de cerrar las casillas, se procede al conteo rápido que tiene como fin único el mostrar un resultado veraz y tan confiable que debe ser muy parecido al resultado oficial. Entre las decisiones más fuertes del conteo rápido es saber cuando se está en condiciones de decir un resultado veraz, el gráfico 4.4.B es un instrumento valioso que ayuda para tomar la decisión.

La encuesta post-electoral nos confirmó lo que la encuesta preelectoral nos dijo, que el voto duro es priísta y que el electorado votó por un cambio en el estado de Nuevo León, también se plantea medir entre la ciudadanía su percepción sobre la transparencia de las elecciones celebradas en el estado de Nuevo León y la mayoría piensa que sí fueron transparentes.

5.2 RESULTADOS COMPARATIVOS

5.2.1 INTRODUCCION

Los cuadros que se presentan en este capítulo, sirven para comparar los resultados que se obtienen en cada uno de los procesos electorales con respecto al resultado oficial, las encuestas que aquí se analizan son la encuesta preelectoral, la encuesta de salida o postvoto, el conteo rápido, la encuesta post-electoral y desde luego los resultados oficiales. Sólo cabe mencionar que en la encuesta de salida, la encuesta preelectoral y post-electoral los que no contestaron, no recuerdo e indecisos se distribuyeron de manera proporcional a los que sí lo hicieron y los nulos se integraron al reactivo otro.

Se muestra un cuadro comparativo del resultado de las encuestas que realizó el Centro de Estudios de Opinión con el resultado oficial para los primeros partidos mayoritarios; también se hace un comparativo sobre el uso de los ponderadores de los estratos en el conteo rápido con el flujo de votación, en otras palabras, se tiene un listado de las personas en lista nominal para la elección del 6 de julio de 1997, la lista estaba actualizada para las elecciones, sin embargo, este modelo supone que los flujos de votación por estrato son los mismos que para el día de la elección, las diferencias se muestran en el cuadro 5-1; asimismo, se hará una comparación sobre los resultados por estrato de cada una de las encuestas con el resultado oficial simultáneamente.

5.2.2 PRESENTACION DE RESULTADOS

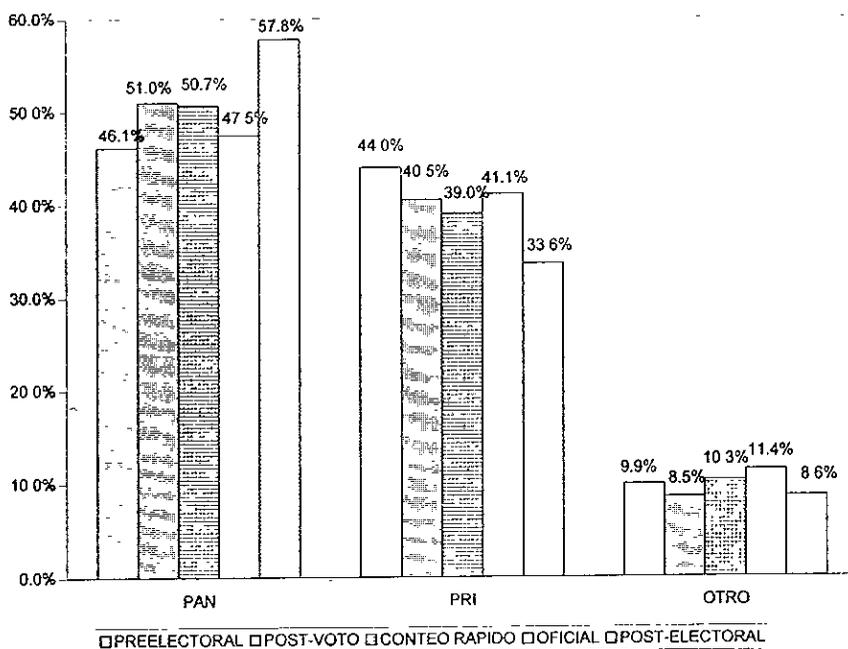
Cuadro 5-1: Comparativo del ponderador utilizado en el conteo rápido con respecto al flujo de votantes por estrato en la elección para gobernador de Nuevo León celebrada el 6 de julio de 1997.

ESTRATO	PONDERADOR	FLUJO DE VOTACION	DIF
1	33.4%	33.5%	-0.1%
2	17.5%	17.4%	0.1%
3	14.0%	14.3%	-0.3%
4	8.5%	8.3%	0.2%
5	3.7%	4.0%	-0.3%
6	5.4%	5.2%	0.2%
7	17.5%	17.3%	0.2%
TOTAL	100%	100%	

El cuadro 5-1 indica que el ponderador que se utilizó el día de la elección fue muy similar al flujo de votación que existió en cada uno de los estratos, por consiguiente, la muestra no tuvo problema alguno en el uso de los ponderadores, el ponderador del estrato i es el porcentaje de personas en lista nominal del estrato con respecto al universo de estudio y el porcentaje de flujo de votación en el estrato i es igual al porcentaje de personas que acudieron a votar en el estrato i con respecto al total de personas que acudieron a votar en el universo de estudio, $i = 1, \dots, 7$, las diferencias de estos resultados son casi nulas.

El gráfico 5.2.A nos presenta los resultados de la encuesta de salida, conteo rápido, post-electoral y el resultado oficial de manera global, se puede visualizar que la más desfasada hacia el PAN es la encuesta post-electoral y la que más subestima al PAN es la encuesta preelectoral, el PRI tuvo una gran baja en la encuesta post-electoral debido a que la gente suele decir que votó por el partido que gana las elecciones con regularidad.

Gráfico 5.2.A Comparativo de la encuesta preelectoral, encuesta de salida, conteo rápido, resultado oficial y encuesta post-electoral en la elección del 6 de julio de 1997 para gobernador del estado de Nuevo León.



Una diferencia importante en la elaboración de la encuesta post-electoral con las otras encuestas, está en que en la primera se definieron doce estratos en la sección 3.5, distintos a los siete estratos definidos en las secciones 3.2, 3.3 y 3.4. Como se explicó en la sección 3.5, la muestra se realizó con asignación proporcional para cada uno de los doce estratos, bajo este diseño redefinimos los estratos para comparar los resultados por estrato de cada una de las encuestas, así, nos queda el cuadro comparativo 5-2 de la distribución de la muestra, la forma en que se reconstruyó el ponderador de la encuesta post-electoral fue dividiendo el total de encuestas que se realizaron en la encuesta post-electoral sobre la base de los estratos definidos en las encuestas preelectoral, de salida y el conteo rápido.

Cuadro 5-2: Comparativo del ponderador de lista nominal con el ponderador hipotético (proporción de entrevistas) de la encuesta post-electoral de la elección para gobernador de Nuevo León celebrada la elección el 6 de julio de 1997.

ESTRATO	PONDERADOR REAL (LISTA NOMINAL)	PONDERADOR ENCUESTA	DIFERENCIA
1	33.4%	33.2%	0.2%
2	17.5%	17.3%	0.2%
3	14.0%	15.2%	-1.2%
4	8.5%	4.8%	3.7%
5	3.7%	3.7%	0.0%
6	5.4%	5.5%	-0.1%
7	17.5%	20.3%	-2.8%
TOTAL	100%	100%	

De esta manera, se hacen los comparativos por estrato, estas gráficas dan la ventaja que se puede visualizar el comportamiento global de cada uno de los estratos para cada una de las cuatro encuestas: preelectoral, encuesta de salida o también llamada postvoto, conteo rápido, post-electoral y el resultado oficial de la elección para gobernador celebrada el día 6 de julio de 1997, los estratos 4 y 7 son los que mayor diferencia tienen, mas los porcentajes de una estratificación y otra no son significativos.

En el gráfico 5.2.B se observa la votación en porcentajes del PAN para cada uno de los estratos, en el gráfico 5.2.C lo mismo que 5.2.B pero para el PRI, en 5.2.C para otros partidos u opciones y en el 5.2.E todos los reactivos en una sola gráfica, de esta manera se visualiza el resumen de los resultados de las tres encuestas en los siete estratos en un solo gráfico.

Gráfico 5.2.B Resultados obtenidos por el PAN por estrato en la encuesta preelectoral, encuesta de salida o postvoto, conteo rápido, resultado oficial y encuesta post-electoral de la elección del 6 de julio de 1997 para gobernador del estado de Nuevo León

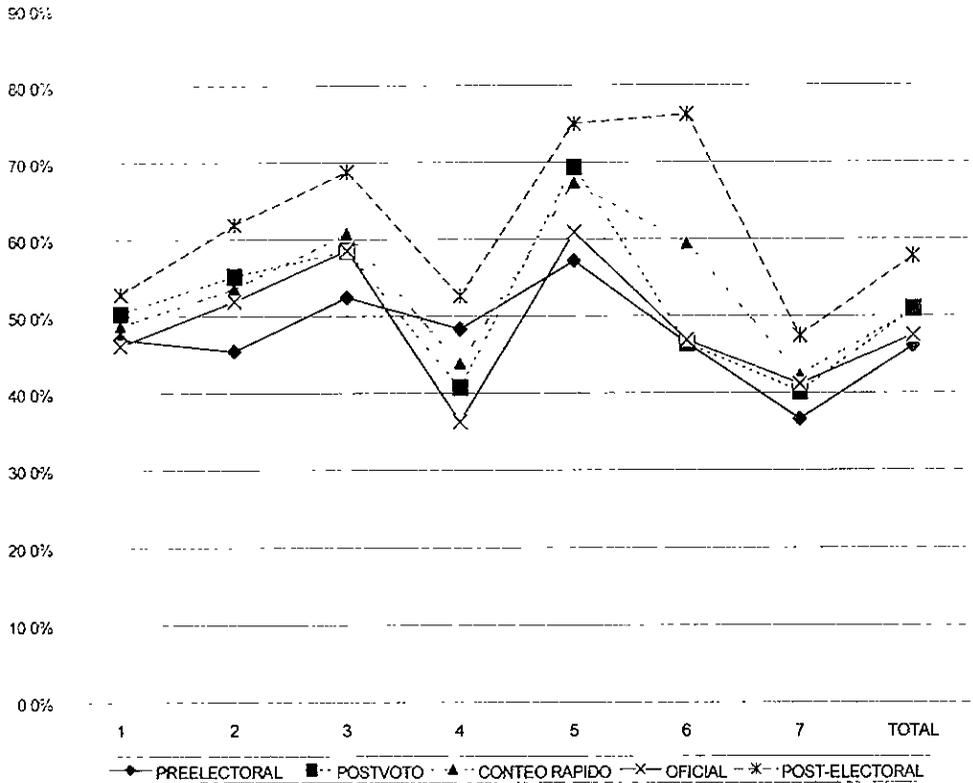


Gráfico 5.2.C Resultados obtenidos por el PRI por estrato en la encuesta preelectoral, encuesta de salida o postvoto, conteo rápido, resultado oficial y encuesta post-electoral de la elección del 6 de julio de 1997 para gobernador del estado de Nuevo León

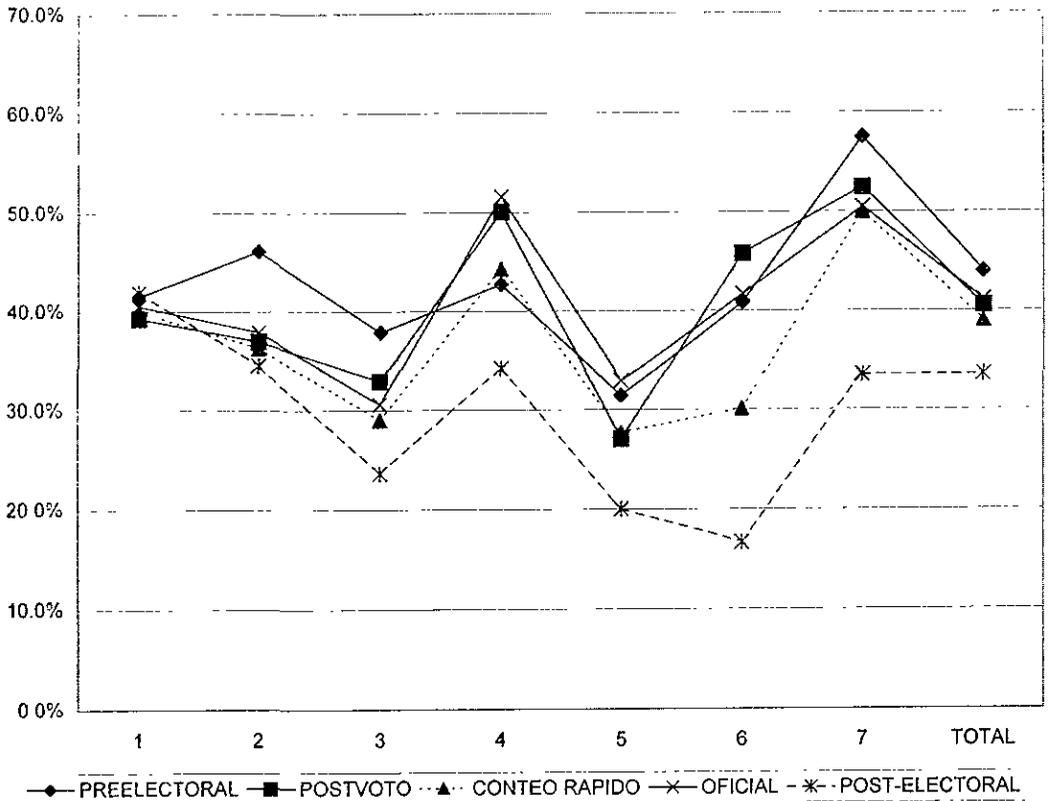


Gráfico 5.2.D Resultados obtenidos por los otros partidos políticos por estrato en la encuesta preelectoral, encuesta de salida o postvoto, conteo rápido, resultado oficial y encuesta post-electoral de la elección del 6 de julio de 1997 para gobernador del estado de Nuevo León

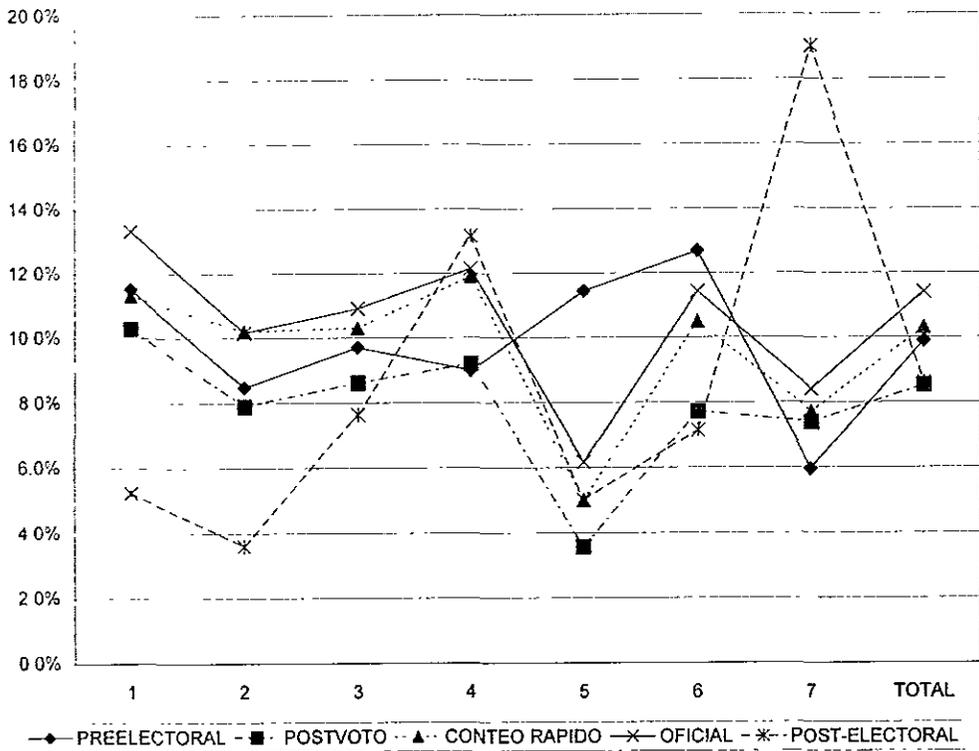
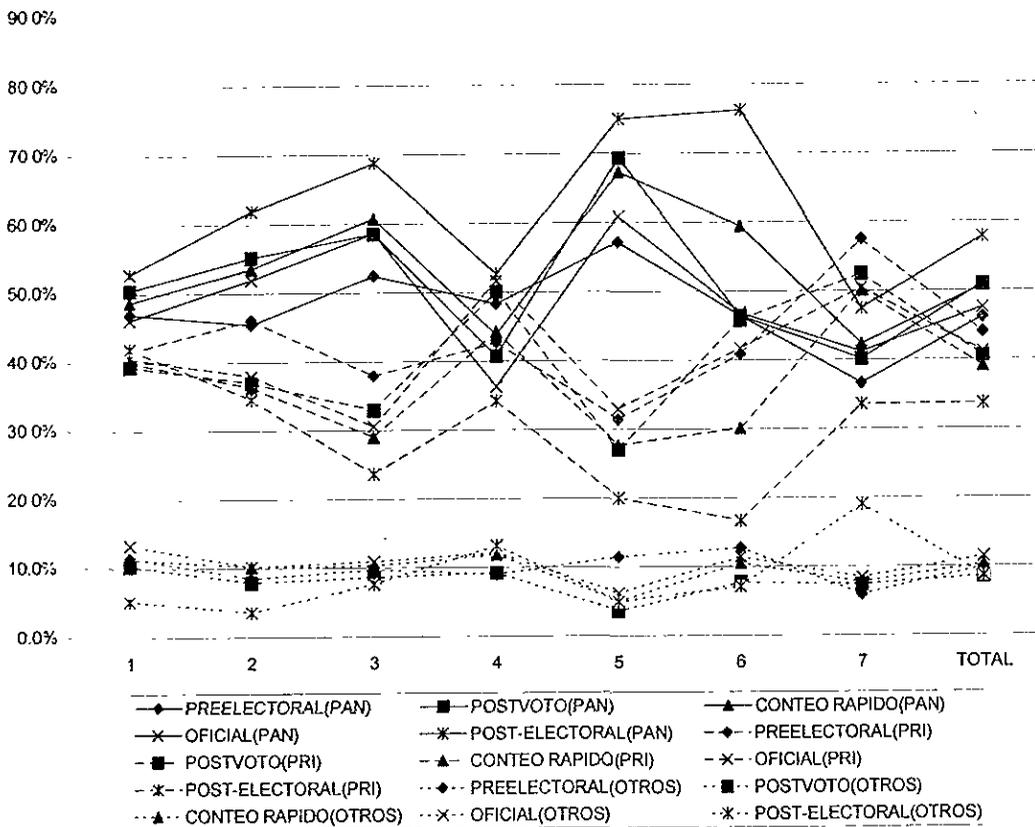


Gráfico 5.2.E Gráfico comparativo de los resultados obtenidos por el PAN, PRI y otros partidos políticos por estrato en la encuesta preelectoral, encuesta de salida o postvoto, conteo rápido, resultado oficial y encuesta post-electoral de la elección del 6 de julio de 1997 para gobernador del estado de Nuevo León



En el gráfico 5.2.B que mide la intención de voto de cada uno de los estratos hacia el PAN, se observa que la encuesta preelectoral tiene 1.4% puntos porcentuales menos que el resultado oficial que es el resultado que más se acerca al resultado oficial, generalmente el resultado que más se acerca al oficial es el conteo rápido y sobre todo si son las mismas secciones electorales, el conteo rápido elimina todas las posibilidades de “mentirosos” y los que “no contestan”, al menos en el caso del PAN el modelo de la encuesta preelectoral de asignar los que no contestaron y los indecisos de manera proporcional entre los que si lo hicieron ha funcionado, pero para el caso de la encuesta post-electoral el PAN se sobrestima en el resultado, esto se debe como ya se mencionó a que la población tiende a decir que votó por el partido que obtuvo el triunfo, sobre todo en los estratos 2 a 7, lo que hace que el resultado se sobrestime el PAN, en el caso de la encuesta postvoto sobrestima ligeramente al PAN, pero el modelo de distribuir a los que no contestaron de manera proporcional entre los que sí lo hicieron ha sido aceptable, para el conteo rápido el resultado es bueno.

Si se analiza el estrato 1 el resultado de la encuesta preelectoral hacia el PAN es ligeramente mayor al resultado oficial, la cantidad de secciones electorales en este estrato es de 61, estas secciones electorales son suficientes para determinar un resultado de un universo homogéneo como lo fue Monterrey, de allí la precisión de este estrato, no solamente en la encuesta preelectoral, también en la encuesta de salida, el conteo rápido y la encuesta post-electoral han sido muy cercanos al resultado oficial.

En el estrato 2 que es el municipio de Guadalupe con 23 secciones electorales el PAN se subestima en la encuesta preelectoral, mas en la encuesta post-electoral se sobrestima, el que el PAN se sobrestime en la encuesta post-electoral es normal debido a lo ya antes explicado, el resultado de postvoto y el conteo rápido son los que más se acercan al resultado oficial.

En el estrato 3 con 20 secciones electores en el municipio de San Nicolás de los Garza, el PAN en la encuesta preelectoral se subestima, la encuesta post-electoral lo sobrestima y la encuesta postvoto y el conteo rápido son los que más se acercan al resultado oficial.

En el estrato 4 con 13 secciones electorales que son los municipios Apodaca y General Escobedo todos los procesos electorales subestimaron al PAN del resultado oficial, el resultado que más se allega fue la encuesta postvoto.

Los estratos 5 y 6 que son San Pedro y Santa Catarina con 5 y 7 secciones electorales respectivamente se decidió partirlos por la diferencia económica que existe entre la población de uno y otro, el estrato 5 tiene una tendencia socioeconómica mayor que la del estrato 6, las diferencias son marcadas y esto fue motivo para estratificarlos, en el estrato 5 que tiene una mejor situación económica el voto hacia el PAN es muy claro, según el resultado oficial que es el estrato que mayor porcentaje de votos obtuvo este partido político, de esa manera se justifica la estratificación a pesar de ser pequeño el estrato, pero en el estrato 7 la encuesta post-electoral tiene una marcada diferencia con respecto al resultado oficial, el resultado de la encuesta post-electoral sobrestima en mucho al PAN, pero cabe recordar que la cantidad de casos no son representativos para cada estrato y este estrato sólo tiene un peso de 3.7% sobre la muestra.

El estrato 7 que tiene 49 secciones en muestra y son los municipios que comprenden el resto del estado, se visitaron un total de 10 de los 44, en la encuesta preelectoral se subestima al resultado oficial, la encuesta postvoto y el conteo rápido se acercan al resultado oficial, la encuesta post-electoral sobrestima al PAN.

La encuesta post-electoral sobrestima en todos los estratos al resultado oficial, el conteo rápido es el proceso que estrato por estrato está más cerca del resultado oficial, pero la encuesta preelectoral se compensó unos estratos subestimados con otros sobrestimados y el resultado global fue el más preciso como se ha mencionado incluso que el conteo rápido.

El gráfico 5.2.C mide la intención de voto de cada uno de los estratos hacia el PRI, la encuesta preelectoral sobrestima ligeramente al PRI, la encuesta postvoto y el conteo rápido están cercanos al resultado oficial y la encuesta post-electoral subestima de manera considerable al PRI, la encuesta preelectoral tiene un comportamiento disperso estrato a estrato pero el que sobrestime en algunos y subestime en otros genera que el resultado de la encuesta preelectoral sea cercano al resultado oficial, el resultado del proceso que más se asemeja al resultado oficial es la encuesta postvoto, pero muy de cerca por el conteo rápido; la encuesta post-electoral subestima al PRI debido a que en los estratos 2 a 6 lo subestima, de esta manera produce el resultado global que subestima al PRI.

En el estrato 1 los resultados de los 4 procesos electorales son muy semejantes, las diferencias entre éstos son pequeñas.

En el estrato 2 la encuesta preelectoral sobrestima al PRI, la encuesta postvoto, post-electoral y el conteo rápido tienen semejanza al resultado oficial.

En el estrato 3 la encuesta preelectoral sobrestima al PRI y la post-electoral lo subestima, el conteo rápido es el proceso más preciso y postvoto es un resultado cercano.

En el estrato 4 el resultado que más se asemeja al resultado oficial es postvoto, pero todos los procesos subestiman al PRI, el conteo rápido y la encuesta preelectoral están a una diferencia de 5% puntos porcentuales, pero la encuesta post-electoral tiene diferencia considerable mayor de 10% puntos porcentuales.

En el estrato 5 se observa que todos los procesos subestiman al PRI, la encuesta preelectoral es el resultado que más se aproxima al resultado oficial, la encuesta post-electoral tiene diferencias considerables mayores a 10% puntos porcentuales, el conteo rápido y la encuesta postvoto tienen una diferencia cercana a los 5% puntos porcentuales, pero cabe recordar que el tamaño de muestra no es representativo para este estrato.

En el estrato 6 la encuesta post-electoral subestima por más de 20% puntos porcentuales al resultado oficial, el conteo rápido subestima también al PRI por más de 10% puntos porcentuales, pero la cantidad de secciones electorales no son representativas para este estrato, la encuesta que más cerca está del resultado es la encuesta preelectoral y postvoto es la única encuesta que sobrestima al PRI.

En el estrato 7 la encuesta sobrestima al resultado oficial, la encuesta postvoto también sobrestima al PRI pero por una diferencia muy pequeña, el conteo rápido es el resultado que más se asemeja al oficial en este estrato, mas la encuesta post-electoral subestima al PRI por más de 15% puntos porcentuales, lo que es una distancia considerable y dada la cantidad de encuestas y peso específico de este estrato.

5.3 CONCLUSIONES Y COMENTARIOS GENERALES

Las herramientas del muestreo en la aplicación de las encuestas electorales son necesarias para que éstas sean confiables y precisas. En México la gente del ámbito político, de los medios de comunicación y el resto de la población cada vez las aceptan más, cabe recordar el gráfico 4.6.P en el que se muestra que la mayoría está de acuerdo en encuestas electorales, ya que dicen favorecer a la democracia y permite la libre expresión de los ciudadanos.

En México las encuestas electorales son indicadores que tienen una gran difusión en la actualidad, sin embargo, esta práctica ya es ampliamente conocida en otros países sobre todo del primer mundo, por ejemplo, en Estados Unidos desde los años treinta en la elección para presidente las encuestas fueron de gran importancia en aquel entonces. Hoy en día en México poco a poco puede verse que este tipo de prácticas son más comunes entre la población mexicana y sobre todo aceptadas por ésta, ya que han venido a fortalecer la democracia, un sondeo de opinión respalda los resultados oficiales ante la sociedad, de esta manera la población mexicana tiene mayor confianza en que los resultados publicados por el órgano electoral correspondiente son verídicos. Por tal motivo es importante hacer notar la responsabilidad profesional y ética que tienen las personas que llevamos a cabo este tipo de estudios, un resultado adverso al oficial puede poner en entredicho al órgano electoral competente, máxime si la publicación de la encuesta es difundida a través de una fuente confiable.

Las encuestas electorales son un gran paso hacia la libre expresión en México, sin embargo, es recomendable que a corto plazo existan normas más concretas y precisas sobre la publicación de las encuestas, debe llegar el momento en que la población sepa distinguir entre las diferentes fuentes confiables, propongo que el error estadístico y el nivel de confianza estén respaldados por un efecto de diseño para conocer que efectivamente la encuesta cumple con los márgenes indicados en el reporte, así mismo, podemos recordar que no existen sanciones para aquellas encuestas publicadas y que no tengan la autorización del instituto electoral. Propongo, además que exista la firma de un responsable de la metodología, éste debe aprobar un examen ante un órgano competente formado por expertos en la materia y avalado por el Instituto Federal Electoral.

Acerca de los resultados que el Centro de Estudios de Opinión obtuvo, se puede ver en el cuadro comparativo 5-3, cada una de las encuestas y el resultado oficial correspondiente; el cuadro presenta el error promedio real de cada uno de los procesos con respecto al resultado oficial:

Cuadro 5-3: Comparativo de los resultados y errores promedio de cada proceso electoral en la elección para gobernador de Nuevo León celebrada el 6 de julio de 1997.

PROCESO	PAN	PRI	OTRO	ERROR
PREELECTORAL	46.1%	44.0%	9.9%	1.9%
POSTVOTO	51.0%	40.5%	8.5%	2.3%
CONTEO RAPIDO	50.7%	39.0%	10.3%	2.1%
OFICIAL	47.5%	41.1%	11.4%	
POST-ELECTORAL	57.8%	33.6%	8.6%	6.9%

El error promedio real se calculó con la suma de las diferencias absolutas de cada estimación por partido político entre tres, según las opciones: PAN, PRI y OTRO.

El proceso más preciso es el conteo rápido por las razones que ya se explicaron, el resultado de la encuesta preelectoral nos demuestra que en la realidad no necesariamente esto se cumple, la encuesta preelectoral del estado de Nuevo León se puede catalogar de excelente, sin embargo la encuesta post-electoral tiene un error de 6.9% debido a la sobrestimación del PAN, pero las causas ya se mencionaron, la encuesta de salida o postvoto tiene un error promedio real de 2.3% por lo que se cataloga como muy buena, debido a que en su momento se tomaron decisiones sobre el acomodo de los que no contestaron y los indecisos, el conteo rápido se cataloga como muy bueno sin dejar de hacer notar que se sobrestimó al PAN y se subestimó al PRI. A pesar de que la encuesta preelectoral y el conteo rápido contaron con las mismas secciones electorales en la muestra. El proceso electoral más cercano a la realidad fue la encuesta preelectoral.

Durante los operativos no existió problema alguno para proceder con el trabajo de campo, las autoridades del estado de Nuevo León dieron todas las facilidades para la realización de estos procesos, los encuestadores hicieron el trabajo que se les encomendó y la capacitación que se les asignó para usar los criterios de selección de los encuestados por género y edad en la encuesta preelectoral, postvoto y post-electoral, también se hicieron de manera adecuada; al encuestador se le dan indicaciones precisas de los perfiles que éstos deben tener y no dejarse llevar por las apariencias de las personas o de

las preferencias de éste. El modelo de la no respuesta de asignarlos de manera proporcional sobre aquellos que sí lo hicieron ha demostrado que para el caso de Nuevo León ha sido bueno.

En un estado como Nuevo León, que es grande territorialmente, sería prácticamente imposible dar un reporte de conteo rápido, veraz y con una aproximación de tres horas posterior al cierre de las casillas. Asimismo, es importante no sólo el marco teórico muestral y la veracidad del levantamiento de encuestas, sino el hecho de considerar a los medios de comunicación. Para el proceso de datos y la generación de tablas de contingencia (cruces) es prioritaria la utilización de la computadora, ya que facilita el diseño de modelos teóricos muestrales con mayor complejidad, aunado al ahorro de tiempo que genera este instrumento. Como conclusión, es importante apearse a la teoría, pero también hacer adaptaciones responsables en la práctica, en el caso de Nuevo León nos demostró lo poderoso de esta herramienta para la toma de decisiones y para conocer la percepción de la población al momento del levantamiento de encuestas con resultados confiables y cercanos a la realidad. Debemos recordar asimismo que no se debe perder de vista el objetivo y el universo del estudio, por ejemplo, el propósito principal de las encuestas en esta tesis fue medir el posicionamiento de los partidos políticos en el estado de Nuevo León, pero hicimos algunos análisis por estrato, por sexo, por edad, etc. Querer obtener información detallada puede distraernos del objetivo principal, además la información en algunas ocasiones no tienen las suficientes encuestas para poder dar mayor soporte, sin embargo, en la práctica lo considero importante para conocer las tendencias tomando las precauciones necesarias.

Considero que no basta saber la teoría del muestreo para hacer encuestas, creo que es importante la experiencia para poder relacionar la teoría con la práctica, también es importante hacer adaptaciones de acuerdo a las circunstancias de cada región geográfica y el objetivo del estudio.

En México existen lugares prácticamente inaccesibles y alejados, por lo que obtener información de éstos es muy costoso y agotador, entrevistar a una persona de estas características sobre su preferencia partidista y temas coyunturales relacionados con el gobierno, este es un ejemplo de la adaptación que se debe realizar cuando se dirige hacia el encuestado, porque de nada serviría un modelo totalmente probabilístico si la población no responde el cuestionario y sobre todo de manera veraz. Aquí se presentó un método eficiente y que al menos para el caso de las elecciones del 6 de julio de 1997

para la gubernatura del estado de Nuevo León arrojó resultados positivos, entrevistar a la gente en la vía pública cuidando que el individuo habite en la unidad de muestreo seleccionada, que para el caso de Nuevo León se tomaron las secciones electorales como puntos de muestreo, por lo que funcionó como se pretendía, esto es una alternativa al muestreo probabilístico.

Las encuestas electorales en México se han convertido en un parámetro de referencia entre las clases políticas, sociales y de los medios de comunicación de nuestro país, éstas seguirán siendo si continúan obteniendo resultados creíbles. En lo personal, vislumbro un futuro prometedor en la industria de las encuestas electorales dada la gran demanda y necesidad de éstas para la mejor toma de decisiones de nuestros gobernantes.

En el año 2000, en México serán las elecciones presidenciales más competidas de la historia, este será un gran reto para las encuestas y se pondrán a prueba los diferentes métodos del levantamiento de las mismas y una vez más el muestreo como herramienta básica.

BIBLIOGRAFIA

Beristain Gil Arturo Fernando,

Análisis estadístico de una encuesta preelectoral, una electoral y una de conteo rápido en el estado de Guanajuato,
Facultad de Ciencias de la Universidad Nacional Autónoma de México.
1997.

Cochran William Gemell,

Sampling Techniques,
Jonh Wiley & Sons, Inc.
3rd edition, 1977.

Cochran William Gemell,

Técnicas de Muestreo,
CECSA, México,
2^a edición, 1984.

Códice90,

Instituto Nacional de Estadística Geografía e Informática (INEGI)
1990

Conteo95,

Instituto Nacional de Estadística Geografía e Informática (INEGI)
1995

Crewe Ivor,

The Opinion Polls: Confidence Restored?
Oxford University Press,
1997.

Curtice John,

Are the opinion polls ready for 1997?
Journal of the Market Research Society Volume 39 Number 2,
April 1997.

Curtice John & Sparrow Nick:

How accurate are traditional quota opinion polls?
Journal of the Market Research Society Volume 39 Number 3,
July 1997.

Distribución Federal del Estado de Nuevo León

Instituto Federal Electoral

Distribución Local del Estado de Nuevo León

Comisión Estatal Electoral del estado de Nuevo León

Haldane, J. B. S.

On a method of estimating frequencies.
Biometrika, 33, 222-225.
1945.

Hansen, M.H., Hurtiwitz, W.N. And Bershad, M,

Measurement errors in censuses and surveys
Bull. Inter. Stat. Inst.
1961.

Hendricks W.A.,

The mathematical theory of sampling
Scarecrow Press, New Brunswick, N. J.
1956.

Kish Leslie,

Survey Sampling
Jonh Wiley & Sons, New York,
1965.

Kish, Leslie,

Def: why, when and how? a review.
ASA Preceedings of Survey Research Methods Section,
Ann arbor: University of Michigan,
1989.

Ley Electoral Del Estado de Nuevo León

13 de Diciembre de 1996

Mendenhall, Scheaffer, Wackerly,
Estadística Matemática con Aplicaciones,
Grupo Iberoamérica
3ª edición, 1986

Neyman, J;

On the different aspects of the representative method:
The method of stratified sampling and the method of purposive selection.
Jour. Roy. Stat. Soc., 97, 558-606
1934.

Reportes Metodológicos, CEO;

Encuesta preelectoral (27 de Junio de 1997)
Encuesta de salida y conteo rápido (Junio de 1997)
Encuesta post-electoral (15 de Julio de 1997)

Centro de Estudios de Opinión de la Universidad de Guadalajara

Stephan, F.F. and Mc Carthy, P.J.,

Sampling Opinions and Analisis of Survey Procedure,
Greenwoos Press Publishers, Westport, Connecticut,
1st edition, 1974.

Sukhatme, P. V.; Sukhatme, B. V.; Sukhatme, S.; Asok, C.,

Sampling Theory of Surveys Applications,
Iowa State University Press
3rd edition, 1984.

Tschuprow, A. A.,

On the mathematical expectation of the moments of frecueny distributions in
the case of correlated observations.
Metron, 2, 461-493, 646-683.
1923.