

13
209



UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS

ANALISIS ESTADISTICO DE UNA ENCUESTA PREELECTORAL, UNA ELECTORAL Y UNA DE CONTEO RAPIDO EN EL ESTADO DE GUANAJUATO.

T E S I S

QUE PARA OBTENER EL TITULO DE

A C T U A R I O

P R E S E N T A :

ARTURO FERNANDO BERISTAIN GIL



DIVISION DE ESTUDIOS PROFESIONALES
DIRECTORA DE TESIS: GUILLERMINA ESLAVA GOMEZ

MEXICO, D.F.

FACULTAD DE CIENCIAS
SECCION ESCOLAR

1997

TESIS CON
FALSA 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.



UNIVERSIDAD NACIONAL
AUTÓNOMA DE
MÉXICO

M. en C. Virginia Abrín Batule
Jefe de la División de Estudios Profesionales de la
Facultad de Ciencias
Presente

Comunicamos a usted que hemos revisado el trabajo de Tesis:
ANÁLISIS ESTADÍSTICO DE UNA ENCUESTA PREELECTORAL,
UNA ELECTORAL Y UNA DE CONTEO RÁPIDO EN EL ESTADO
DE GUANAJUATO.

realizado por Arturo Fernando Beristáin Gil.

con número de cuenta 8752256-9 , pasante de la carrera de ACTUARIO

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

Director de Tesis

Propietario

Dra. Guillermina Eslava Gómez.

Propietario

M. en C. José Antonio Flores Díaz.

Propietario

Mat. José Luis Castrejón Caballero.

Suplente

M. en C. Salvador Zamora Muñoz.

Suplente

Act. Javier Ibarra Piña.

Consejo Departamental de Matemáticas
Act. Agustín Román Aguilar

A MIS PADRES
C.P. J. FELIPE BERISTÁIN JUÁREZ
MARA A. GIL HERNÁNDEZ
EN AGRADECIMIENTO A SU PACIENCIA
Y A SU APOYO INCONDICIONAL

A DAHLIA Y ARIADNE
POR TODO LO QUE APRENDÍ
DE ELLAS COMO HERMANAS
Y COMO MUJERES

A MIS AMIGOS:

RODRIGO, POR SU IMPULSO Y SUS CONSEJOS

WENCESLAO, POR SUS ENSEÑANZAS Y TRUCOS
COMPUTACIONALES

JORGE Y RODULFO, POR NO DEJARME SOLO EN
CIENCIAS

RUBEN, POR SU AYUDA

A la mujer con la que compartiré mi vida y los recuerdos de
este trabajo, esperando comprenda porque no esta su
nombre

Umberto Eco sostiene que "Es de mal gusto dar las gracias al ponente¹. Si os ha ayudado no ha hecho más que cumplir con su obligación".² En lo personal no cumulo con este pensamiento; tal vez en Europa así sea el estilo, pero en Latinoamérica la mentalidad es distinta, por lo que no seguiré esta recomendación y, con todo respeto para Umberto Eco, expreso el siguiente agradecimiento.

A LA DRA. GUILLERMINA ESLAVA GOMEZ -EMINENCIA EN ESTADÍSTICA APLICADA- POR SU ERUDITA DIRECCIÓN Y COMO MUESTRA DE AGRADECIMIENTO POR HABER CREÍDO EN MÍ Y PERMITIRME PARTICIPAR EN ALGUNAS DE SUS INVESTIGACIONES Y TRABAJOS

¹ director de tesis

² Umberto Eco, *Como se hace una Tesis*, versión castellana, España, Gedisa, S.A., 1992 pág. 218.

GRACIAS A

M. en C. JOSÉ ANTONIO FLORES DÍAZ

Mat. JOSÉ LUÍS CASTREJÓN CABALLERO

M. en C. JOSÉ SALVADOR ZAMORA MUÑOZ

Act. JAVIER IBARRA PIÑA

Catedráticos de estadística y sus diversas ramas, por cuidar que llegue con bien a esta meta e indicarme el camino hacia las siguientes

**A LA UNIVERSIDAD NACIONAL AUTÓNOMA DE
MÉXICO QUE COLOCÓ EN MI DESTINO A LA
FACULTAD DE CIENCIAS Y ME DIO LA
OPORTUNIDAD DE OBTENER ESTE
PRIMER GRADO ACADÉMICO**

**AL CENTRO DE ESTUDIOS DE OPINIÓN DE LA
UNIVERSIDAD DE GUADALAJARA PORQUE SU
VALIOSO MATERIAL HIZO POSIBLE ESTE
TRABAJO**

**AL INSTITUTO ESTATAL ELECTORAL DE GUANAJUATO
Y A LA 2ª CIRCUNSCRIPCIÓN DEL INSTITUTO FEDERAL
ELECTORAL POR LA INFORMACIÓN PROPORCIONADA**

PREFACIÓN

El propósito de este análisis es describir tres mecanismos de información que tienen un alto grado de relación con la celebración de comicios electorales, el primero se aplica por lo menos una vez en un período cercano a la jornada, cuyo nombre es *Encuesta Preelectoral*; el segundo, por lo general es de aplicación única y se lleva a cabo inmediatamente después de que un elector ha emitido su voto, conocido como *Encuesta de Salida*; el tercero y último, se planea para ser ejecutado justo en el momento en que se cierra oficialmente una casilla electoral y se llama *Conteo Rápido*. El lugar donde se realizaron los ejercicios citados fue el estado de Guanajuato en sus elecciones extraordinarias para gobernador en mayo 28 de 1995. Las encuestas fueron diseñadas por el Centro de Estudios de Opinión (CEO), organismo dependiente de la Universidad de Guadalajara.

El análisis se desarrollará en seis capítulos: el capítulo uno contiene información geográfica, estadística y sociodemográfica del estado de Guanajuato tal como: población total, distribución geográfica, composición por edad y sexo, población con derecho a voto etc., tomada en su mayoría del "Perfil sociodemográfico del estado de Guanajuato", publicado por el Instituto Nacional de Estadística Geografía e Informática (INEGI). Además de los datos anteriores, se presentan las divisiones electorales por distritos en sus modalidades federal y estatal, la primera realizada por el Instituto Federal Electoral (IFE) y la segunda por el Instituto Electoral del Estado de Guanajuato (IEE Guanajuato). También contiene un espacio dedicado al muestreo aleatorio simple.

El capítulo dos contiene el soporte teórico de algunos de los diversos diseños muestrales que se utilizaron en dichos comicios. Estos diseños son el muestreo estratificado bietápico, el muestreo bietápico y el muestreo por conglomerados, aplicados a la encuesta preelectoral, la encuesta de salida y al conteo rápido respectivamente. Como estos diseños son realmente combinaciones de técnicas de muestreo, se tratan independientemente cada una y también algunas de sus variantes o procesos auxiliares como el muestreo por cuotas.

El capítulo tres es la encuesta preelectoral, cuya estructura está dividida prácticamente en análisis sociodemográfico particular de la información de la encuesta y en el análisis de preferencias, sin faltar la descripción del operativo de campo ni la logística utilizada para la misma.

Los capítulos cuatro y cinco son la encuesta de salida y el conteo rápido respectivamente, cuya estructura desde un punto de vista práctico es la misma; en ambos se presentan la descripción de los ejercicios, su operatividad, el tamaño de muestra y , lo más importante, los resultados de cada uno.

Los análisis de los capítulos tres al cinco son de tipo descriptivo más que inferencial y, aún cuando se hacen comparaciones entre categorías, se debe recordar su propia naturaleza. Los análisis son descriptivos porque el objetivo de la tesis es exhibir la aplicación de los diseños de muestreo en mecanismos de medición de tendencias electorales que, por ser de reciente divulgación en nuestro país, permiten hacer una retroalimentación para modelarlos de acuerdo a las necesidades de las características del proceso electoral o tomar una decisión para sustituirlos por diseños más eficientes y fáciles de adaptar.

Se debe tener claro y estar consciente de la cantidad de teoría, investigación y desarrollo que están detrás de los diseños de muestreo utilizados, cuyos resultados en este trabajo se utilizan sin demostrar rigurosamente los teoremas que los sustentan.

En el capítulo seis se resumen los resultados obtenidos de cada encuesta para que después de realizar una comparación entre ellas, se puedan establecer las conclusiones pertinentes de la investigación.

Gracias a la amabilidad del IEE Guanajuato fue posible realizar la comparación de los estimadores obtenidos por cada ejercicio con los resultados oficiales en forma desagregada de las elecciones. Las comparaciones se hacen en el capítulo correspondiente a cada ejercicio.

Por último, se presenta un anexo que contiene: 1) los errores más comunes ajenos al muestreo que surgen en encuestas de salida y conteos rápidos; 2) la tabla resumen que engloba todos los municipios y la información recabada por las tres investigaciones agrupadas por estratos.

INTRODUCCIÓN	1
CAPITULO 1 Preliminares	3
1.1 Perfil sociodemográfico del estado de Guanajuato	3
1.1. 1 Población total	3
1.1. 2 Distribución geográfica	3
1.1. 3 Crecimiento de la población	4
1.1. 4 Composición por edad y sexo	4
1.1. 5 Migración	4
1.1. 6 Alfabetismo	4
1.1. 7 Asistencia escolar	5
1.1. 8 Fecundidad	5
1.1. 9 División electoral federal	6
1.1.10 División electoral estatal	7
1.2 Muestreo aleatorio simple (MAS)	12
1.2.1 Motivación	12
1.2.2 Teoría del muestreo	13
1.2.3 Notación del MAS	13
1.2.4 Tamaño de muestra	15
1.2.5 Estimadores de proporciones y de sus varianzas	16
CAPITULO 2 Otros Diseños de Muestreo y Algunos Procesos Auxiliares	19
2.1 Muestreo estratificado	19
2.1.1 Descripción	19
2.1.2 Notación	20
2.2 Muestreo por conglomerados	22
2.2.1 Descripción	22
2.2.2 Notación	24
2.3 Submuestreo o Muestreo bietápico	25
2.3.1 Descripción	25
2.3.2 Notación	25
2.4 Submuestreo estratificado	29
2.4.1 Descripción	29
2.4.2 Notación	29

2.5 Muestreo por cuotas	31
2.5.1 Descripción	31
CAPITULO 3 Encuesta Preelectoral	35
3.1 Introducción	35
3.1.1 Descripción	35
3.1.2 Diseño muestral	36
3.1.3 Aplicación del muestreo estratificado bietápico	36
3.1.4 Operativo de campo	39
3.2 Resultados	41
3.2.1 Tamaño de la muestra	41
3.2.2 Perfil sociodemográfico de la muestra	41
3.2.3 Asociación entre algunas Variables Manejadas en la Encuesta Preelectoral	44
3.2.3A Variables electorales vs. variables sociodemográficas	45
3.2.3B Preferencias partidarias vs. variables sociodemográficas	68
3.2.4 Preferencias por partido en cada estrato	79
3.2.5 Preferencia partidista	84
3.2.6 Repulsión partidista	90
3.2.7 Comparativo preferencia vs. rechazo	94
CAPITULO 4 Encuesta de Salida	98
4.1 Introducción	98
4.1.1 Descripción	98
4.1.2 Diseño Muestral	98
4.1.3 Aplicación del muestreo bietápico	103
4.1.4 Operativo de campo	104
4.2 Resultados	104
4.2.1 Tamaño de muestra	104
4.2.2 Preferencias	104
4.2.3 Preferencias por partido en cada estrato	105
CAPITULO 5 Conteo Rápido	107
5.1 Introducción	107

5.1.1	Descripción	107
5.1.2	Diseño Muestral	107
5.1.3	Aplicación del muestreo por conglomerados	108
5.1.4	Operativo de campo	113
5.2	Resultados	113
5.2.1	Tamaño de muestra	113
5.2.2	Tendencias	114
5.2.3	Tendencias por cada estrato	116
CAPITULO 6 Conclusiones		118
Anexo		128
Bibliografía		131

INTRODUCCION

En la actualidad vivimos en un mundo cuyo movimiento está regido por la información, es tal nuestra dependencia de ella, que algunas veces no es posible ejecutar una idea si no tenemos por lo menos un indicio de lo que vamos a hacer. Un ejemplo para ilustrar es el siguiente, suponga que hará alguna tarea doméstica, digamos un cambio de tubería del agua, si lo hace uno mismo, de inmediato se piensa en algún conocido que ya lo haya realizado para que nos refiera su experiencia, de esta forma uno está pidiendo información antes de efectuar cualquier acción; luego se pregunta si tuvo éxito en su cometido, esto es, una vez obtenida la información, se requiere de una **muestra** de eficiencia o más directamente de un resultado para decidir si uno mismo realiza el trabajo o bien, contrata a alguien. La intención del ejemplo es explicar que muchas acciones quizá sin importancia en otro contexto están involucradas con la necesidad de obtener una muestra previa si es que existe, para poder tomar la mejor decisión.

Una herramienta muy poderosa para obtener información es el muestreo. En el contexto matemático, herramienta y técnica frecuentemente suelen ser sinónimos; para decirlo con elegancia, el muestreo es una técnica muy útil en muchos tipos de investigación.

Las investigaciones basadas en encuestas por muestreo se dividen en dos grandes grupos, el primero lo representan las encuestas a nivel nacional, que por sus altos costos sólo las pueden financiar organismos especializados pertenecientes al gobierno federal de un país, u organismos internacionales como la O.N.U., entre otros; el otro grupo es relativamente de menor escala, pero aun así se trata de gobiernos estatales y municipales o de grandes industrias transnacionales.

Dentro del primer grupo podemos contar, planes nacionales tales como alimentación, producción, desarrollo y en general áreas que estén ligadas a la economía nacional.

Para el segundo grupo, casi son los mismos rubros referidos a estados o municipios; para los grandes complejos industriales los estudios son utilizados con fines de expansión y optimización, tanto de producción como de comercialización, e incluso para el área comercial y ventas la investigación adquiere su propio nombre: estudio de mercado, en donde puede estar involucrado el lanzamiento de un nuevo producto o la mejora de otro ya existente.

En el presente trabajo, nos referiremos a un estudio similar a las investigaciones de mercado, pero cuya conceptualización correcta es estudio de opinión.

Quando se escucha el término "estudio de opinión", es común relacionarlo de inmediato con la popularidad de una persona que es capaz de acceder a cierta candidatura; pero este tipo de investigación va más allá, porque un estudio de opinión también es útil para saber qué tanta aceptación tiene un producto después de haberlo lanzado al el mercado, o bien, si algún otro producto ya clásico en el mercado, aún sigue manteniendo su hegemonía y qué tan susceptible sería un cambio después de permanecer en un cierto nivel; en fin, hay ejemplos más cotidianos como los sondeos de nivel de audiencia para programas radiofónicos o televisivos. Después de aclarar que un estudio de opinión no sólo sirve para cuestiones políticas, es necesario recordar que este trabajo analiza las posibilidades de todos los candidatos a la primera magistratura del estado de Guanajuato, a través de información registrada por la encuesta preelectoral.

El estudio se basó en la encuesta preelectoral a los comicios de Guanajuato, para analizar, además del grado de popularidad de los candidatos contendientes en ese momento (14 de mayo de 1995), algunos factores relacionados con el ambiente previo a las elecciones, tales como: qué proporción del electorado contaba con la credencial de elector, si conocía a que partido pertenecía el gobernador saliente, si conocía la fecha de las elecciones, si conocía los nombres de los candidatos contendientes y a que partido pertenecen, etc. .

Este estudio también cuenta con el análisis estadístico descriptivo de dos ejercicios relativamente novedosos en nuestro país, que son: la encuesta de salida y el conteo rápido.

La finalidad de este trabajo es presentar los diseños muestrales y su aplicación, así como reflejar la aplicación del muestreo y el análisis descriptivo de la información obtenida, más que los resultados de los estudios de opinión.

CAPITULO 1 Preliminares

Este capítulo se divide en dos apartados: el 1.1 de carácter sociodemográfico y de información electoral cuyos datos demográficos se tomaron del "Perfil Sociodemográfico XI Censo General de Población y Vivienda, 1990; volumen Guanajuato", publicado por el INEGI. La información electoral es de los municipios que integran cada distrito tanto en el sistema federal (IFE) como en el estatal (IEE Guanajuato). Se incluyen dos mapas con las divisiones distritales citadas anteriormente y otro con la división municipal del estado realizada por el INEGI, finalmente este apartado se cierra con una tabla que contiene las claves, nombres, población total, población con derecho a voto (de 18 años cumplidos en adelante) de cada municipio, así como las proporciones de ambas poblaciones sobre el total del estado y el total de cada municipio. El apartado 1.2 describe brevemente la motivación del uso del muestreo aleatorio simple.

1.1 Perfil Sociodemográfico del estado de Guanajuato

1.1.1 Población Total

La población de 3,982,593 habitantes registrada por el "XI Censo General de Población y Vivienda, 1990" en el estado de Guanajuato, comparada con la de 1980 (3,006,110 hab.), arroja una tasa de crecimiento media anual del 2.9% en esa década, mientras que la anual a nivel nacional es de el 2.05%.

1.1.2 Distribución Geográfica

León es el municipio que concentra la mayor población del estado con el 21.8%; le siguen Irapuato con 9.1% y Celaya con el 7.8%. El resto de los municipios presenta un porcentaje de población que se encuentra en un intervalo entre el 0.1% y 5.1%. La tabla 1 presenta la distribución de la población total, y la mayor de 18 años, los pesos sobre sus respectivas poblaciones y el peso de la población potencial con derecho a voto sobre la población total en cada municipio. La información fue extraída del código 90 del INEGI

El censo registró un total de 6,617 localidades, observándose que el 63.4% de la población reside en localidades con 2,500 habitantes o más, mientras que el restante 36.6% lo hace en localidades con menos de 2,500 habitantes. El INEGI posee su propia definición de población urbana, describiéndola como aquella que reside en localidades mayores o iguales a 2,500 personas o que son cabecera municipal.

También se observa una tendencia de concentración en localidades grandes, pues el 50.9% de la población reside en localidades con más de 20,000 habitantes que constituyen el 0.3% del total de localidades; por otro lado el 14.2% de la población estatal se distribuye en el 84.6% de pequeñas localidades con menos de 500 habitantes y esto indica que también existe dispersión de la población en localidades pequeñas.

1.1.3 Crecimiento de la Población

Las tasas de crecimiento de los municipios se encuentran en el rango del 0.3% al 5.4%, aunque la mayoría sobrepasa el 2.0%.

1.1.4 Composición por Edad y Sexo

La población guanajuatense tiene una estructura por edad "joven", porque los menores de 15 años constituyen al 41.2% del total, mientras que los de 65 años y más sólo participa con un 4.3% del total; haciendo una comparación con el total del país tenemos que, la población menor a 15 años de la República Mexicana representa el 36% del total; puede decirse que Guanajuato tiene una población "joven" mientras que la nación está transitando de una población joven a una "entrada en años". La "edad mediana" es de 18 años con lo que se reafirma la aseveración. Cabe señalar que el grupo 0 - 4 años, que es la base de la pirámide, registra una proporción menor a la del grupo siguiente (5 - 9 años), lo que puede significar descensos de la fecundidad y por lo tanto, el inicio de un proceso de "envejecimiento" de la población. Por sexo, el 48.4% son hombres y el 51.6 mujeres.

Con la ayuda de estos datos también podemos saber que la población de 18 años y más constituye el 50.07% del total de la población del estado, es decir, esta es la población potencial con derecho a voto.

1.1.5 Migración

En materia de migración conocemos que el 7.8% de la población censada no era nativa de la entidad. Por otro lado, en 1985 la población de 5 años y más que no residía en Guanajuato representa un 3.2%.

Por municipios, Tarandacuao es el municipio con mayor porcentaje de no nativos en 1990, con el 13.1%, seguido de Celaya con 12% y Purísima del Rincón con un 11.5%. En lo que se refiere a proporciones de residentes fuera de la entidad en 1985, destacan los municipios de Tarandacuao con el 7.1% y Celaya con un 5.2%. Cabe señalar que Guanajuato tiene el quinto lugar a nivel nacional en cuanto a migración hacia los Estados Unidos de Norteamérica¹ con un 7%, en el período de 1980 a 1990; del primero al cuarto lo ocupan Chihuahua, Jalisco, Distrito Federal y Michoacán, respectivamente.

1.1.6 Alfabetismo

En el estado, el 85.8% de la población de 6 a 14 años sabe leer y escribir, respecto a la población de 15 años y más, el 83.2 % son alfabetas. Al comparar con los censos anteriores, se observa que el porcentaje de analfabetas en la entidad ha disminuido significativamente en los últimos 20 años, al pasar del 37.5 % en 1970 a 16.5 % en 1990.

¹ Tomado de "Programa Nacional de Población 1995-2000" publicado en el Diario Oficial de la Federación el 7 de agosto de 1995.

Por municipios se notan variaciones importantes, por ejemplo, Xichú con el 37.2 %, y Tierra Blanca con el 32.6 % de analfabetas; existen otros tres municipios que se encuentran en este rango de analfabetismo, en contraste los porcentajes más bajos de analfabetas los presenta la capital del estado (Guanajuato) con 11.0 % y León con el 11.1 %.

En relación con las diferencias por sexo, del total de analfabetas, el 62.4 % son mujeres y el 37.6 % hombres.

1.1.7 Asistencia Escolar

En el estado de Guanajuato el 78.8 % de la población de 5 a 14 años asiste a la escuela; se observa también que la asistencia aumenta conforme avanza la edad hasta los 9 años, donde se presenta un máximo del 92.8 %; a partir de la siguiente edad se inicia un descenso de la asistencia.

Con respecto a diferencias por sexo tenemos que no presentan mayores diferencias hasta los 11 años; sin embargo, para las edades siguientes la asistencia de los hombres es mayor que la de las mujeres.

1.1.8 Fecundidad

La comparación de los promedios de hijos nacidos por mujer en 1990, 1980 y 1970 muestra que de 3.4 hijos nacidos en 1970 disminuyó a 3.2 en 1980 y luego a 2.8 para 1990. Por municipios, este promedio en 1990 presenta variaciones de consideración: Atarjea tiene el promedio más alto con 3.8 hijos por mujer, un hijo más que el promedio estatal; mientras que el promedio más bajo se encuentra en Moroleón, con 2.5 hijos por mujer.

Según el Diario Oficial de la Federación del 7 de agosto de 1995 la tasa a nivel nacional en el período 1987 - 1991 es de 3.5 hijos por mujer; la tasa global de fecundidad de Guanajuato es de 3.9 hijos. La tasa más alta a nivel estatal la poseen Chiapas y Oaxaca con 4.6 hijos y la más baja el Distrito Federal con 2.2 hijos. Guanajuato ocupa el lugar número 10.

1.1.9 División electoral federal

Fuente: IFE, 2ª circunscripción.

<p>DISTRITO I</p> <p>Guanajuato Romita Silao</p>	<p>DISTRITO VI</p> <p>Salamanca Pueblo Nuevo Abasolo Huanimaro</p>	<p>DISTRITO IX</p> <p>Allende Atarjea Dr. Mora San José Iturbide San Luis de la Paz Santa Catarina Sta. Cruz de J. Rosas Tierra Blanca Victoria Xichú</p>	<p>DISTRITO XII</p> <p>Valle de Santiago Moraleón Santiago Maravatio Uriangato Yuriria</p>
<p>DISTRITO II</p> <p>León Noreste</p>	<p>DISTRITO VII</p> <p>Acámbaro Apaseo el Alto Apaseo el Grande Coroneo Tarandacua Jerécuaro</p>	<p>DISTRITO X</p> <p>Dolores Hidalgo Ocampo San Diego de la Unión San Felipe</p>	<p>DISTRITO XIII</p> <p>Salvatierra Cortázar Jaral del Progreso Villagrán Tarimoro</p>
<p>DISTRITO III</p> <p>León Sur</p>	<p>DISTRITO VIII</p> <p>Celaya Comanfort</p>	<p>DISTRITO XI</p> <p>León Noroeste</p>	
<p>DISTRITO IV</p> <p>Irapuato</p>			
<p>DISTRITO V</p> <p>Cuéramaro Manuel Doblado Pénjamo Purísima del Rincón San Fco. del Rincón</p>			

1.1.10 División electoral estatal

Fuente: IEE Guanajuato.

DISTRITO I Guanajuato Silao	DISTRITO VII Irapuato Oeste Romita	DISTRITO XII Moroleón Salvatierra Santiago Maravatío	DISTRITO XVI Allende Comonfort San José Iturbide
DISTRITO II León Noreste	DISTRITO VIII Irapuato Este Cuéramaro	Uriangato Yuriria	DISTRITO XVII Atarjea Dr. Mora San Luis de la Paz Santa Catarina Tierra Blanca Victoria Xichú
DISTRITO III León Sur	DISTRITO IX Sta. Cruz de J. Rosas	DISTRITO XIII Acámbaro Coroneo Jerécuaro Tarandacua	
DISTRITO IV León Noroeste	Salamanca Este Villagrán		
DISTRITO V Manuel Doblado Purísima del Rincón San Fco. del Rincón	DISTRITO X Salamanca Oeste Pueblo Nuevo	DISTRITO XIV Apaseo el Alto Apaseo el Grande Celaya Este Tarimoro	DISTRITO XVIII Dolores Hidalgo Ocampo San Diego de la Unión San Felipe
DISTRITO VI Abasolo Huanímaro Pénjamo	DISTRITO XI Jaral del Progreso Valle de Santiago	DISTRITO XV Celaya Oeste Cortázar	

Comparando ambas divisiones, se observa que los tres municipios más importantes aparte de León, (Celaya, Irapuato y Salamanca); fueron hechas por el instituto local de tal forma, que cada una de las partes pertenecen a distintos distritos, además, es curioso que el distrito VIII de la división estatal, no está integrado en una sola región, es decir, existen dos regiones separadas geográficamente pero están catalogadas con el mismo número de distrito. Esto se notará más fácil en los mapas de ambas divisiones de las figuras 1 y 2. Por otra parte, sabemos que el IIE de Guanajuato está haciendo una reestructuración sobre la división distrital y que se verá incrementado el número de distritos electorales, de dieciocho a veintidós. También el IFE anunció que de trece elevará a quince los distritos electorales en Guanajuato.

GUANAJUATO

DISTRITOS ELECTORALES FEDERALES

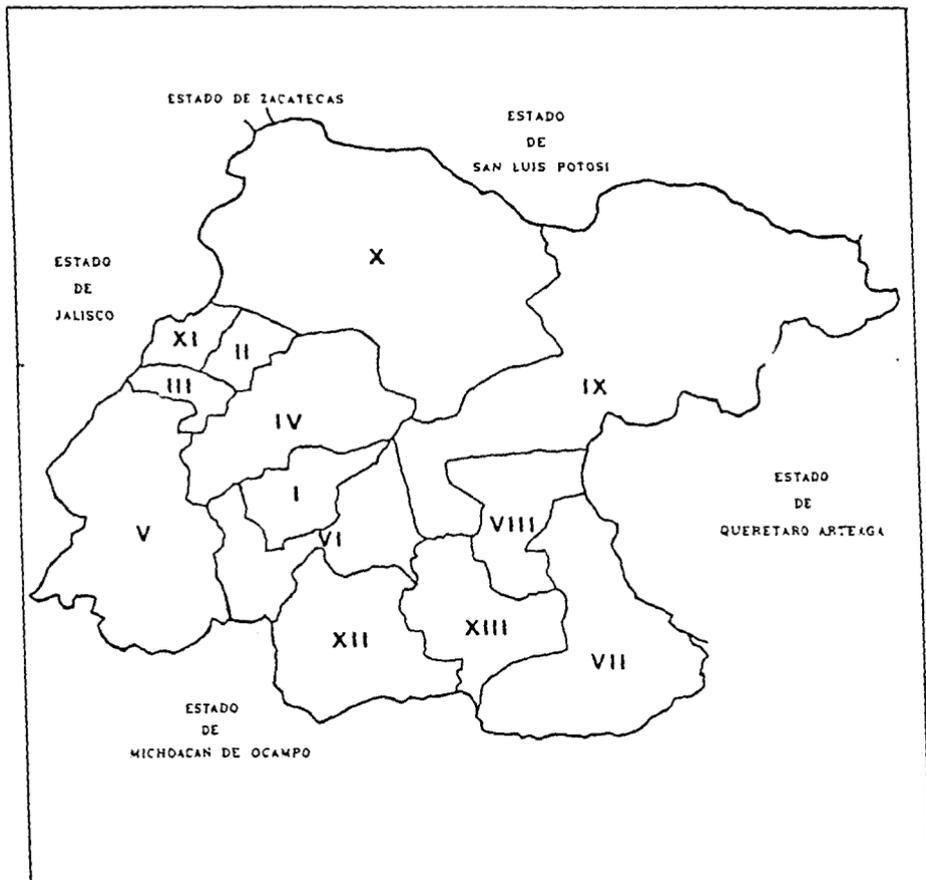


Figura 1. Distritos electorales diseñados por el IFE

GUANAJUATO

DISTRITOS ELECTORALES ESTATALES

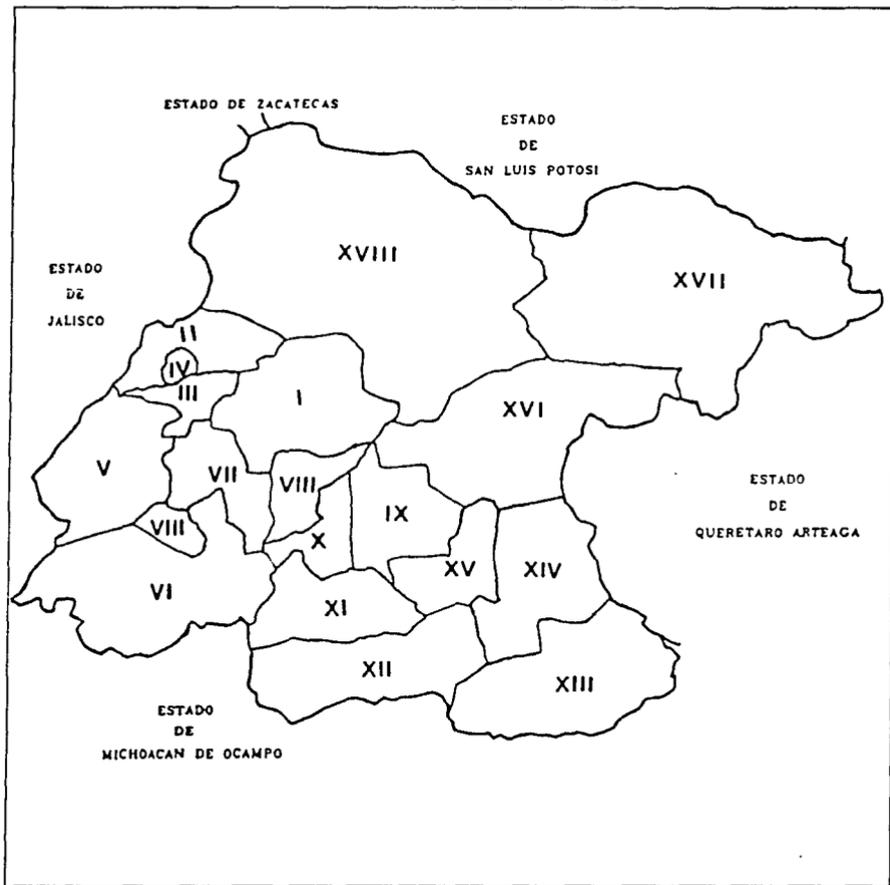


Figura 2. Distritos electorales diseñados por el IEE de Guanajuato

GUANAJUATO

DIVISIÓN MUNICIPAL

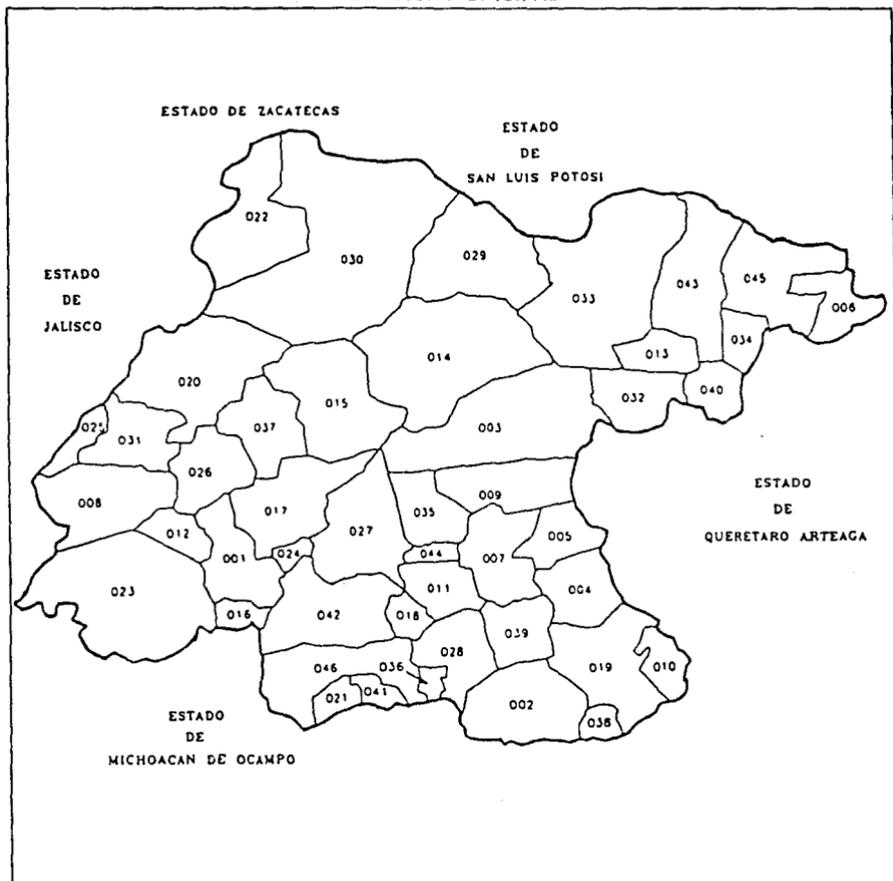


Figura 3. División Municipal del estado de Guanajuato según perfil sociodemográfico INEGI 1990 (claves de los municipios en la tabla 1)

GUANAJUATO
DISTRIBUCIÓN DE POBLACIÓN TOTAL
Y POBLACIÓN POTENCIAL CON DERECHO A VOTO
POR MUNICIPIOS

Clave del municipio según INEGI	NOMBRE del MUNICIPIO	Población total por municipio	Población mayor de 18 años	Peso relativo municipios por su pob total	Peso relativo pob > de 18 sobre pob tot del municipio	Peso relativo municipios por su pob mayor de 18
1	ABA SOLO	70938	34200	1.78	48.34	1.72
2	ACAMBARO	112450	57558	2.82	51.19	2.89
3	ALLENDE	110602	53180	2.78	48.04	2.67
4	APASEO EL ALTO	48455	22771	1.22	46.90	1.14
5	APASEO EL GRANDE	61594	29364	1.55	47.67	1.47
6	ATARJEJA	5200	2561	0.13	49.25	0.13
7	CELAYA	310569	163012	7.80	52.49	8.17
8	MANUEL DOBLADO	37484	17807	0.94	47.51	0.89
9	COMONFORT	56592	27014	1.42	47.73	1.35
10	CORONEO	9435	4784	0.24	50.70	0.24
11	CORTAZAR	74383	38565	1.87	51.85	1.93
12	CUERAMARO	23109	11254	0.58	48.70	0.56
13	DOCTOR MORA	16814	7927	0.42	47.15	0.40
14	DOLORES HIDALGO	104712	47210	2.63	45.09	2.37
15	GUANAJUATO	119170	59342	2.99	49.80	2.98
16	HUANIMARO	18300	8813	0.46	48.00	0.44
17	IRAPUATO	362915	187427	9.11	51.64	9.40
18	JARAL DEL PROGRESO	29764	15719	0.75	52.81	0.79
19	JERÉCUARO	51954	23967	1.30	46.13	1.20
20	LEON	867920	446299	21.79	51.42	22.38
21	MOROLEON	48191	25623	1.21	53.17	1.28
22	OCAMPO	19727	8740	0.50	44.30	0.44
23	PENJAMO	137842	68040	3.46	49.36	3.41
24	PUEBLO NUEVO	10507	5391	0.27	50.87	0.27
25	PURISIMA DEL RINCON	30433	14306	0.76	47.01	0.72
26	ROMITA	44545	20703	1.12	46.48	1.04
27	SALAMANCA	204311	109438	5.13	53.56	5.49
28	SALVATIERRA	97509	51531	2.45	52.80	2.58
29	SAN DIEGO DE LA UNIÓN	31360	14270	0.79	45.50	0.72
30	SAN FELIPE	87216	39914	2.19	45.76	2.00
31	SAN FRANCISCO DEL RINCON	83601	40596	2.10	46.56	2.04
32	SAN JOSÉ ITURBIDE	42681	20465	1.07	47.95	1.03
33	SAN LUIS DE LA PAZ	78504	36683	1.97	46.73	1.84
34	SANTA CATARINA	3982	2023	0.10	50.80	0.10
35	SANTA CRUZ DE JUVENTINO ROSAS	56166	27912	1.41	49.70	1.40
36	SANTIAGO MARAVATIO	9483	4830	0.24	50.93	0.24
37	SILAÓ	115130	54658	2.89	47.48	2.74
38	TARANDACUAO	12402	6208	0.31	50.06	0.31
39	TARIMURO	38594	18767	0.97	48.63	0.94
40	TIERRA BLANCA	13121	6185	0.33	47.14	0.31
41	URIANGATO	46710	23958	1.17	51.29	1.20
42	VALLE DE SANTIAGO	132023	65573	3.32	49.67	3.29
43	VICTORIA	18324	9084	0.46	49.57	0.46
44	VILLAGRAN	38680	19589	0.97	50.64	0.98
45	XICHU	11614	5508	0.29	47.43	0.28
46	YURIRIA	77247	35302	1.94	45.70	1.77
	GUANAJUATO	3982593	1994161	100	50.07	100

Tabla 1. Distribución de población total y con derecho a voto por municipio. Fuente: Códice 90 del INEGI

1.2 Muestreo aleatorio simple (MAS)

1.2.1 Motivación

Ya se ilustró que en algunos casos se requiere de una muestra antes de emprender algo, para explicar mejor lo que es el Muestreo Aleatorio Simple (MAS) se presenta el siguiente ejemplo, supongamos que un vendedor novato de pavos, realizará su primera venta importante que consta de 500 piezas; el comprador condiciona que los pavos pesen como mínimo 3 kilos, nuestro empresario acude con su proveedor y hace la compra, sin prever la condición del cliente, en camino a cerrar el negocio con éste, piensa que el comprador no se va a lanzar a la aventura de pesar 500 animales, además de que confía en la experiencia y buena fe de su proveedor.

Una vez que llega con el comprador, éste le pide que descargue el producto porque llevará a cabo un experimento, el cual consiste en la elección de 12 pavos al azar, nuestro personaje piensa de inmediato "¿qué desconfianza!", pero se trata de una venta importante y confía en que el producto cumple con la exigencias. Después de pesar las piezas, es notificado de que 7, pesan en promedio de 2.7 kgs. y, además, enfáticamente le hacen saber que están congelados, es decir, que el peso del agua congelada está influyendo sobre el peso real del pavo; por lo tanto, la transacción se puede cancelar.

Después de este relato, las preguntas son: ¿Qué tan representativo es sacar 12 pavos de una totalidad de 500? ¿Cuál sería la decisión si hubieran sido menos de 7 pavos los que no cumplieron con el peso?

Las respuestas se darán más adelante; lo sobresaliente del ejemplo es que puede ser real y la persona que aplicó la técnica de elegir al azar cierto número de productos, ni siquiera sabe que efectuó una técnica muy cercana al MAS, sin embargo, la realización del experimento lo hizo dudar sobre llevar a cabo o no la operación.

El método más simple de muestreo probabilístico es el MAS. A continuación se describe su proceso: se seleccionan r unidades en un conjunto cuya totalidad es N , esto quiere decir que si se etiquetan todos los elementos del conjunto se pueden usar etiquetas numeradas desde 1 hasta N , esto implica que el conjunto es finito y que se conoce el número total de sus elementos, para elegir los r elementos o unidades debe haber seguridad de que tengan la misma probabilidad de ser extraídos. Cuando este confirmado lo anterior, se sacan una por una de las r piezas seleccionadas al azar o aleatoriamente.

Aplicando el procedimiento al ejemplo, esto es lo que debió hacer en un principio el distribuidor con el mayorista: a) asignar un número de 1 a 500 a cada pavo, b) acomodar los pavos de tal forma que se puedan sacar sin ningún problema, acomodándolos en una hilera, o tal vez en un arreglo matricial de 50 columnas por 10 filas, de manera que sea fácil llegar a alguno en específico y poder retirarlo de su ubicación, y c) con una calculadora con función generadora de números aleatorios (que generalmente dan un número de tres decimales entre 0 y 1), se realizará el siguiente proceso 12 veces. ¿Por qué 12 veces?, porque se elegirán 12 pavos.

Se genera dicho número que se denota como *RND*. Como son 500 pavos, se multiplica *RND* por 500 y el resultado obtenido se redondea; por último se localiza el pavo que está etiquetado con dicho número redondeado y se extrae para llevarlo a la báscula.

Se propuso el método de la máquina porque en la actualidad es más sencillo tener una calculadora con la citada función, que tener una tabla de números aleatorios, pero si se poseen las tablas y también el proceso de cómo utilizarlas, la mecánica es exactamente igual.

Tal vez la descripción aplicada al ejemplo pueda ser artificial en lo que respecta al MAS, pero reiteramos que la situación del ejemplo le puede suceder a cualquiera.

1.2.2 Teoría del muestreo

Entrando a terrenos más específicos, los estudios por muestreo se hacen en la práctica a base de encuestas que se aplican sobre la población a investigar.

"La palabra *población* se emplea para denominar el conjunto del que se elige la muestra..." (Cochran, 1990, pág. 24)¹

Por lo tanto, desde este momento las llamaremos encuestas por muestreo. No es sencillo elaborar un encuesta por muestreo, porque se deben considerar varios aspectos que en algún momento no tienen importancia, pero acumulados constituyen un gran problema para el desarrollo operativo de la encuesta, entre los más típicos están el lenguaje, la desconfianza, la transportación hasta el operativo de campo, y situaciones tan extrañas que solamente permaneciendo en el trabajo de campo es como se puede adquirir experiencia al respecto, lo curioso es que en varias etapas de las encuestas el muestreo no interviene, sin embargo, estas etapas son muy importantes para la realización del estudio.

La teoría del muestreo, cuyo principio fundamental es tener la máxima precisión a un costo mínimo, toca un punto que no se había mencionado anteriormente, se trata del costo; esta es la razón por la que las encuestas por muestreo se convierten en un negocio. Es tan noble la teoría del muestreo que resulta ser autoanalítica, porque se utiliza para predecir la precisión y el costo que esperamos. Si se sabe cual es el costo, es posible calcular la ganancia potencial. Por otro lado, la precisión no se puede predecir, debido a que no se conoce el verdadero error de una estimación; pero, la precisión se puede juzgar al examinar la distribución de frecuencia, suponiendo que el proceso se aplica en repetidas ocasiones a la misma población.

1.2.3 Notación del MAS

Dentro de los objetivos de la encuesta generalmente están las propiedades que deseamos medir, que se llaman atributos. Los atributos son proporcionados por la unidad de muestreo elegida y, prácticamente, al mismo tiempo en que se proporcionan se registran en la encuesta. Existen distintas notaciones, aquí se adoptará la utilizada por Sukhatme et al (1984, p 25 y 26).

¹ Cochran W., *Técnicas de Muestreo*, 1990 México versión en español (1ª edición, 8ª reimpresión CECSA) pp. 24

Hay algunas variantes que se deben aclarar en cuanto a la notación utilizada, el autor denota a S^2 como los cuadrados medios de la población con denominador $N-1$, es decir, multiplica $\frac{1}{N-1}$ por la suma de los cuadrados de las diferencias de las observaciones con respecto al valor medio poblacional, esto se aprecia mejor en la fórmula (1.2); en versiones traducidas al español de otros títulos sobre la materia e incluso en obras escritas originalmente en español, S^2 es una modificación de la **varianza** poblacional (denotada por σ^2 y con denominador N), algunos de estos libros justifican esta modificación argumentando que es más sencillo trabajar con esta ligera diferencia, y otros dicen que la utilizarán para posteriormente demostrar que su respectiva estimación, es un estimador insesgado de la varianza de la población. Sukhatme denota explícitamente a σ^2 como la varianza poblacional pero, en términos de S^2 y, aunque lo presenta con un factor hasta cierto punto extraño, no es difícil comprobar que este factor se convierte en $\frac{1}{N}$ como se ve en (1.4). En (1.6) sucede lo mismo, en obras en español s^2 es la varianza muestral que es precisamente el estimador insesgado de S^2 y, por lo tanto, de σ^2 .

Sea Y el atributo que estará sujeto a medición y registro, denotaremos con Y_i el valor del atributo asociado con la unidad U_i , donde $i = 1, \dots, N$

$$\bar{Y} = \frac{1}{N} \sum_{i=1}^N Y_i, \text{ la media de la población} \quad (1.1)$$

$$S^2 = \frac{1}{(N-1)} \sum_{i=1}^N (Y_i - \bar{Y})^2 = \frac{1}{N-1} \left[\sum_{i=1}^N Y_i^2 - N\bar{Y}^2 \right], \text{ cuadrados medios de la población} \quad (1.2)$$

$$T(Y) = \sum_{i=1}^N Y_i, \text{ la población total} \quad (1.3)$$

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (Y_i - \bar{Y})^2 = \frac{N-1}{N} S^2, \text{ la varianza poblacional} \quad (1.4)$$

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i, \text{ la media muestral} \quad (1.5)$$

$$s^2 = \frac{1}{(n-1)} \sum_{i=1}^n (y_i - \bar{y})^2, \text{ cuadrados medios de la muestra} \quad (1.6)$$

$$V(\bar{y}) = \left(\frac{N-n}{N} \right) \frac{S^2}{n} = (1-f) \frac{S^2}{n}, \text{ varianza de la media muestral} \quad (1.7)$$

donde $f = \frac{n}{N}$ se conoce como la fracción de muestreo

$$\hat{V}(\bar{y}) = (1-f) \frac{s^2}{n}, \text{ estimador de la varianza de la media muestral} \quad (1.8)$$

$$SE(\bar{y}) = \sigma_{\bar{y}} = \sqrt{\frac{N-n}{N} \frac{S^2}{n}} = \sqrt{(1-f)} \frac{S}{\sqrt{n}}, \text{ error estándar de } \bar{y} \quad (1.9)$$

1.2.4 Tamaño de la muestra

El tamaño de la muestra es literalmente el principio del diseño de la encuesta por muestreo, porque es uno de los primeros pasos después de haberse discutido el diseño, el costo y la operación del método muestral cuya importancia es básica dentro del desarrollo de la investigación.

Como todas las etapas en el diseño de una encuesta por muestreo, la estimación del tamaño de la muestra tiene su propia metodología; si se elige una muestra muy grande se corre el riesgo de caer en gastos superfluos y desperdicio de recursos, por el contrario, una muestra pequeña decaería en eficiencia, lo que provocará que los resultados de la encuesta sean débiles. Frecuentemente, no hay suficiente información para seleccionar o estimar el tamaño de la muestra, por lo tanto, no siempre es sencillo tomar una decisión al respecto.

Para aclarar cualquier duda sobre el tamaño de muestra, se regresará al ejemplo de los pavos, suponga que el vendedor nos pide que le demos una sugerencia sobre el número de pavos que debe seleccionar de los 500 que va a adquirir, de tal suerte que tenga una muestra representativa y confiable para poder decidir si el total de los pavos cumple con los requerimientos que el cliente le pide. Es decir, pide definir el tamaño de la muestra; hasta este momento sólo tenemos N que es la población total, el resto de la información nos la hará saber el vendedor que, de acuerdo a su experiencia, algunos de estos datos los puede estimar.

Enseguida le pedimos nos indique qué es lo que necesita; nos pide que la muestra le pueda dar un parámetro de decisión sobre la compra de los animales, y que sería muy ventajoso saber a través de la muestra si por lo menos el 70% de los pavos pesan como mínimo 2.8 kgs.. Hasta este punto está muy claro lo que desea el comerciante; le contestamos que es posible calcular lo que pide pero, de acuerdo a sus requerimientos, la muestra será grande y por más que intentemos reducirla, tal vez le parezca aún grande con respecto a la población total; además es necesario hacerle un par de advertencias, una es que existe un riesgo de variación, por lo que nos debe dar una tolerancia y la otra es que no podemos garantizar quedar dentro del rango de variación tolerado. Una vez establecido lo anterior nos dice que está dispuesto a tolerar un máximo del 10%, es decir, que el 60% pese 2.8 kgs. al que denotaremos con d ; por nuestra parte, le informamos que el riesgo de que se sobrepase el error (denotado por α) será de digamos un 5%, que visto en términos de confiabilidad estamos hablando de un 95%.

Ya se puede estimar un número que por el momento es lo único pero, como se mencionó anteriormente, aún es posible reducirlo; se denota como n_0 cuya expresión es la siguiente

$$n_0 = \frac{t^2 pq}{d^2} \quad (1.10)$$

donde t es la abscisa de la curva normal que corta un área de α en las colas de la distribución, p es el porcentaje de muestra y q es $1-p$, por lo tanto tenemos:

$$t = 1.96 \quad d = 0.10 \quad p = 0.70$$

$$n_0 = \frac{(1.96)^2 (0.7)(0.3)}{(0.1)^2} = \frac{0.8067}{0.01} = 80.67 \cong 81$$

por lo tanto el vendedor debe extraer una muestra de 81 pavos, pero, este número le puede parecer grande y, tomando en cuenta que le cuestan los procesos de seleccionar y pesar cada pieza, podemos ofrecerle un refinamiento de este número gracias a que la población es finita. Por lo tanto, un tamaño de muestra definitivo se puede obtener con la siguiente expresión

$$n = \frac{n_0}{1 + \frac{(n_0-1)}{N}} \quad (1.11)$$

sustituyendo $n_0 = 81$ y $N = 500$ se obtiene

$$n = \frac{81}{1 + \frac{(81-1)}{500}} = \frac{81}{1 + \frac{80}{500}} = \frac{81}{1.16} = 69.83 \cong 70$$

Tal vez no sea una diferencia muy grande pero, si al comerciante le causa un importante ahorro en sus costos de operación, es obvio que seleccionará 70 pavos. Por otra parte, si él considera que le da lo mismo 81 pavos que 70, debemos sugerirle seleccione el mayor número de pavos.

1.2.5 Estimadores de proporciones y de sus varianzas

La idea de una proporción es, generalmente, una parte de algo, pero en muestreo se identifica con la parte que posee cierta característica de una población o bien que pertenecen a una clasificación definida. Por ejemplo, en la población de una escuela, es común preguntar por la proporción de alumnos que no serán promovidos al siguiente periodo escolar, o bien cual es el porcentaje de personas que no nacieron en el lugar donde residen actualmente. La clasificación generalmente se diseña con un tipo de pregunta que solo acepta como respuesta "SI" o "NO"; en la mayoría de los casos "SI" se introduce como 1 en las fórmulas para obtener los resultados pertinentes y el "NO", o el complemento, como 0.

NOTACIÓN

Suponiendo que todas las unidades de la población se pueden clasificar en dos clases mutuamente excluyentes, es decir, que solamente hay dos tipos y ninguna puede ser de ambos tipos, se denotará con C y C' cada tipo, entonces:

Al número de unidades de la población que caen en C las denotan con A ; y al número de unidades de la muestra con la misma característica, con a .

Para las proporciones, se denota a las unidades de la población con $P = \frac{A}{N}$; y a las de la muestra con $p = \frac{a}{n}$.

La estimación para P es p y la estimación muestral de A es Np ó $\frac{Na}{n}$.

En estimación de proporciones y , tomará el valor de 1 si se encuentra en el tipo C y 0 si está en C' por lo que

$$Y = \sum_{i=1}^N y_i = A \quad (1.12)$$

$$\bar{Y} = \frac{\sum_{i=1}^N y_i}{N} = \frac{A}{N} = P \quad (1.13)$$

Para la muestra.

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{n} = \frac{a}{n} = p \quad (1.14)$$

Para expresar S^2 y s^2 en términos de P y p veamos primero que

$$\sum_{i=1}^N y_i^2 = A = NP \quad (1.15)$$

$$\sum_{i=1}^n y_i^2 = a = np \quad (1.16)$$

entonces

$$S^2 = \frac{\sum_{i=1}^N (y_i - \bar{Y})^2}{N-1} = \frac{\sum_{i=1}^N y_i^2 - N\bar{Y}^2}{N-1} \quad (1.17)$$

utilizando (1.13) y (1.15) que son P y NP respectivamente, tenemos

$$\begin{aligned} S^2 &= \frac{1}{N-1}(NP - NP^2) \\ &= \frac{1}{N-1}NP(1-P) = \frac{N}{N-1}P(1-P) \end{aligned} \quad (1.17a)$$

sea $Q = 1 - P$ entonces

$$S^2 = \frac{N}{N-1}PQ \quad (1.18)$$

haciendo exactamente lo mismo para s^2 tenemos que

$$s^2 = \frac{n}{n-1}pq \quad (1.19)$$

Para un muestreo aleatorio simple de las unidades clasificadas tenemos:

$p = \frac{a}{n}$ es un estimador insesgado de la proporción de población $P = \frac{A}{N}$

La varianza de p es

$$\begin{aligned} \text{Var}(p) &= E(p - P)^2 = \frac{S^2}{n} \left(\frac{N-n}{N} \right) \\ &= \frac{NPQ}{n(N-1)} \left(\frac{N-n}{N} \right) \end{aligned}$$

$$\text{Var}(p) = \frac{PQ}{n} \left(\frac{N-n}{N-1} \right) \quad (1.20)$$

La varianza de la estimación del número total de unidades en la clase C es

$$\text{Var}(\hat{A}) = \frac{N^2 PQ}{n} \left(\frac{N-n}{N-1} \right) \quad (1.21)$$

El estimador insesgado de la varianza para una muestra es (Cochran, 1990, pág. 47)

$$v(\bar{y}) = \frac{s^2}{n} \frac{N-n}{N} \quad (1.22)$$

cambiando p por \bar{y} y sustituyendo (1.19) en (1.22) tenemos

$$\begin{aligned} v(p) &= \frac{\frac{n}{n-1} p q}{n} \frac{N-n}{N} \\ &= \frac{n p q}{n(n-1)} \frac{N-n}{N} \\ v(p) &= \frac{N-n}{N(n-1)} p q \end{aligned} \quad (1.23)$$

introduciendo s_p^2 entonces

$$v(p) = s_p^2 = \frac{N-n}{N(n-1)} p q \quad (1.24)$$

El estimador de la varianza de \hat{A} es el número total estimado de unidades en la clase C en la población, cuya expresión es

$$v(\hat{A}) = s_A^2 = \frac{N(N-n)}{n-1} p q \quad (1.25)$$

El intervalo de confianza por **aproximación normal** para la varianza estimada de p obtenida en (1.24) es

$$p \pm \left[t \sqrt{\left(\frac{1-n}{N} \right) \frac{pq}{(n-1)} + \frac{1}{2n}} \right] \quad (1.26)$$

Donde t es el desvío normal correspondiente a la probabilidad de confianza y $\frac{1}{2n}$ es una corrección para mejorar la aproximación; por esto se remarca que es una aproximación normal, que hasta cierto punto es limitada, porque depende de las cantidades de n , p , N y de las cantidades asociadas a los límites de confianza. Incluso existen tablas de reglas para trabajar y decidir en qué casos es posible utilizar la aproximación normal. También existen otros métodos calificados como exactos y otro tipo de aproximaciones, sobretodo si p es pequeña (ver Cochran, 1990, p. p. 87-90).

CAPITULO 2 Otros Diseños de Muestreo y Algunos Procesos Auxiliares

Este capítulo está dedicado a mostrar el aspecto teórico de las técnicas utilizadas por cada una de las encuestas aplicadas antes y durante los comicios. También se mencionan los métodos que sirven de auxilio a estos diseños. El muestreo estratificado bietápico que fue utilizado para la encuesta preelectoral combina el muestreo de dos etapas con el muestreo estratificado, esto indica que los diseños no son exclusivos, por el contrario, en general se combinan para robustecerse mutuamente y ofrecer mayor eficiencia en la estimación.

Los apartados del capítulo describen el muestreo estratificado, el muestreo por conglomerados, el submuestreo o muestreo en etapas y el submuestreo estratificado. Cada apartado se divide en descripción y notación; en esta última se presentan también los estimadores correspondientes a cada diseño. Finalmente, se explica brevemente un método de selección que está relacionado con el muestreo estratificado, se trata del muestreo por cuotas.

2.1 Muestreo Estratificado

2.1.1 Descripción

La precisión de un estimador no solo depende del tamaño de la muestra y de la fracción de muestreo, sino que también de la heterogeneidad que haya entre las unidades de la población.

Por otro lado, si se va incrementando el tamaño de la muestra, existen otras formas de estimar la media de la población con alto grado de precisión, que consiste en dividirla en grupos, en los que se espera que sean más homogéneos que la totalidad de la población y en los que además se extraerá una muestra de tamaño específico. Los grupos en los que se divide la población se llaman estratos y al proceso de aplicar cualquier diseño de muestreo en cada uno de los estratos se le conoce como muestreo estratificado.

Por ejemplo para estimar el ingreso promedio familiar, sería lógico agrupar a las familias en dos o más grupos o estratos de acuerdo a sus erogaciones por concepto de renta. Así las familias pertenecientes a cada estrato, estarían más homogéneamente agrupadas con respecto a su ingreso que si se les considerara dentro de la totalidad de la población.

De esta forma el ingreso estimado por familia basado en un muestreo estratificado tendría altas posibilidades de ser más preciso que el muestreo aleatorio simple, considerando que ambos se aplican a la misma población y con el mismo tamaño de muestra.

Lo que se busca con el muestreo estratificado es que los elementos que se encuentren en un estrato determinado tengan características similares entre sí, es decir, una varianza pequeña; y de manera simultánea deben tener una diferencia notable al compararlos con elementos que pertenezcan a otro estrato. Estableciendo lo mismo en términos de estratos, se trata de que haya homogeneidad dentro de cada estrato, pero, entre estratos debe haber heterogeneidad notable.

En el muestreo aleatorio simple no hay garantía de que los segmentos de la población tengan cierto grado de representatividad en la muestra. En contraparte, con el muestreo estratificado es

posible sacar una muestra con el grado de representatividad que se desee e incluso realizar estimaciones diferentes en cada una de las subpoblaciones. En este caso se denomina *dominio de estudio* a cada subpoblación. Estas son algunas de las ventajas que ofrece la estratificación y del porqué es tan frecuentemente utilizada para el diseño de encuestas de muestreo. Evidentemente el muestreo estratificado tiene como supuesto que el tamaño de los estratos es conocido y que existe la disponibilidad de un marco de selección para extraer una muestra de cada estrato.

Así pues, el muestreo estratificado aparte de aumentar el grado de precisión de la estimación de la media poblacional, también es común usarlo para obtener la estimación del total de la población, la media de todos los estratos y transformar ciertos estratos que constituyen la población en dominios de estudio.

La estratificación es muy recomendable para planes a gran escala tales como planes nacionales, porque no es necesario tener datos exactos de lo que se va a investigar sino más bien, alguna guía principal en la que se pueda dividir la población sujeta a estudio, como por ejemplo la división política de una región, la clasificación por comunidades, o alguna otra clasificación censal, el objetivo es tratar de buscar un estrato conocido.

Por último, en casos en que no se dispone del tiempo suficiente para planear, el muestreo estratificado brinda un gran apoyo para realizar un diseño eficiente en forma rápida.

2.1.2 Notación

Partiendo de la suposición de que una población de tamaño N se divide en L estratos, y que además, dentro de cada estrato se realiza una muestra con algún otro diseño; para este caso se aplica MAS. Entonces para el i -ésimo estrato con $i=1, 2, \dots, L$. Se denota por:

$N = \sum_{i=1}^L N_i$, como el número total de estratos de la población

N_i , como el número total de unidades en el estrato i

n_i , como el tamaño de muestra dentro del estrato

$f_i = \frac{n_i}{N_i}$, como la fracción de muestreo (2.1)

Y_{ij} , como el valor del atributo bajo estudio para la j -ésima unidad, con $j=1, \dots, N_i$

$W_i = \frac{N_i}{N}$, ponderación del estrato (nótese que el divisor de la ponderación es N) (2.2)

$\bar{Y}_i = \frac{1}{N_i} \sum_{j=1}^{N_i} Y_{ij}$, la media del estrato (sobre el número de unidades N_i) (2.3)

$\bar{y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij}$, como la media de la muestra (sobre la unidades n_i) (2.4)

$$S_i^2 = \frac{1}{(N_i - 1)} \sum_{j=1}^{N_i} (Y_{ij} - \bar{Y}_i)^2, \text{ como los cuadrados medios del estrato} \quad (2.5)$$

$$s_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2, \text{ como los cuadrados medios en la muestra} \quad (2.6)$$

Para evitar confusiones, no se debe olvidar que la notación anterior está referida al estrato i , es decir, la muestra se está haciendo sólo dentro del estrato i , por eso no se puede hablar de parámetros poblacionales totales. Pero, si se puede hablar en términos de parámetros poblacionales para el estrato i .

ESTIMACIONES

La media de la población en términos de la notación anterior es

$$\bar{Y} = \frac{\sum_{i=1}^I \sum_{j=1}^{N_i} Y_{ij}}{N} = \frac{\sum_{i=1}^I N_i \bar{Y}_i}{N} = \sum_{i=1}^I W_i \bar{Y}_i \quad (2.7)$$

por (2.5) y (2.2)

como se está aplicando MAS en el estrato, \bar{y}_i es un estimador de \bar{Y}_i ; al introducir el estimador en (2.7) obtenemos la estimación para la media poblacional cuya expresión es

$$\bar{y}_{or} = \sum_{i=1}^I W_i \bar{y}_i \quad (2.8)$$

Existen algunas condiciones y resultados importantes detrás de la varianza de \bar{y}_{or} . Las condiciones son que \bar{y} sea un estimador insesgado de \bar{Y} en todos y cada uno de los estratos, y que la extracción de las muestras en los distintos estratos sean independientes (esto implica que no necesariamente se debe extraer la muestra de cada estrato con muestreo aleatorio simple). Los resultados son que \bar{y}_{or} es una función lineal de las medias \bar{y}_i con ponderaciones fijas W_i , y usando el resultado de la varianza de una función lineal, tal como la suma de variables aleatorias independientes, tenemos que la varianza está dada por

$$V(\bar{y}_{or}) = \sum_{i=1}^I W_i^2 V(\bar{y}_i) \quad (2.9)$$

donde $V(\bar{y}_i)$ es la varianza de las estimaciones de las medias \bar{Y}_i en el estrato correspondiente.

Hasta este punto, las expresiones anteriores son aplicables al *muestreo estratificado*, esto es muy importante tenerlo en mente porque, como ya se mencionó antes, si se aplica un MAS en cada estrato, el diseño se convierte en un muestreo aleatorio estratificado. Además, por la manera en que se manejó el proceso surge un resultado al que no se le dió la atención necesaria, pero que es importante, se trata del hecho de que \bar{y}_{or} es un estimador insesgado de la media poblacional \bar{Y} (Sukhatme et al, 1984 pág. 107).

A continuación se obtendrá dentro del marco de el muestreo aleatorio estratificado la varianza de \bar{y}_{st} .

Tomando al estrato en forma individual y utilizando las condiciones anteriores y (1.7) del capítulo anterior, se tiene

$$Var(\bar{y}_i) = \frac{S_i^2}{n_i} \left(\frac{N_i - n_i}{N_i} \right) \quad (2.10)$$

sustituyendo (2.10) en (2.9)

$$\begin{aligned} V(\bar{y}_{st}) &= \sum_{i=1}^L W_i^2 \left(\frac{S_i^2}{n_i} \right) \left(\frac{N_i - n_i}{N_i} \right) \\ &= \sum_{i=1}^L W_i^2 \left(\frac{S_i^2}{n_i} \right) \left(1 - \frac{n_i}{N_i} \right) \\ V(\bar{y}_{st}) &= \sum_{i=1}^L W_i^2 \frac{S_i^2}{n_i} \underbrace{\left(1 - f_i \right)}_{\text{por (2.1)}} \end{aligned} \quad (2.11)$$

la expresión anterior constituye un resultado exclusivo del muestreo aleatorio estratificado.

Introduciendo s_i^2 , que es el estimador de S_i^2 en (2.11), y transformando el término $(1 - f_i)$ tenemos que

$$\hat{V}(\bar{y}_{st}) = \sum_{i=1}^L W_i^2 \left(\frac{N_i - n_i}{N_i} \right) \frac{s_i^2}{n_i} \quad (2.12)$$

es una estimación insesgada de la varianza de \bar{y}_{st} . (Sukhatme et al, 1984 pág. 107)

Para los intervalos de confianza se dan

$$\bar{y}_{st} \pm t s(\bar{y}_{st}) \text{ para la media poblacional} \quad (2.13)$$

$$N \bar{y}_{st} \pm t N s(\bar{y}_{st}) \text{ para la población total} \quad (2.14)$$

donde t es el acostumbrado desvío normal correspondiente a la probabilidad de confianza. Estas expresiones solo funcionan bajo el supuesto de que \bar{y}_{st} se encuentre normalmente distribuida y que $s(\bar{y}_{st})$ esté plenamente definida. Para casos especiales que no cumplan con lo anterior, como por ejemplo un número reducido de grados de libertad; existen otros métodos (ver Cochran, 1990, pp. 132-133).

2.2 Muestreo por Conglomerados

2.2.1 Descripción

Este proceso muestral también presupone la división de la población en un número finito de unidades, las cuales se conocen como unidades de muestreo. La unidad de muestreo más pequeña en la que se puede dividir una población es el elemento de la población, y a los grupos de elementos se les llama conglomerados. Cuando una unidad de muestreo es un conglomerado, el proceso muestral es llamado muestreo por conglomerados.

En muestreo por conglomerados se busca que al dividir la población en subpoblaciones, haya alta heterogeneidad dentro de cada conglomerado pero, entre conglomerados deben haber bastantes similitudes u homogeneidad.

Si una gran región contiene población que va a ser encuestada, es posible subdividirla en pequeños segmentos de área, y a cada elemento de la población lo asociamos con uno y solamente uno de estos segmentos de área, el proceso es llamado alternativamente muestreo por área. Cuando una lista de elementos no está disponible o no es confiable es obvio que no es conveniente utilizar sus elementos como unidades de muestreo. El método de muestreo por conglomerados o por área es idóneo en estos casos. En una ciudad podría ser posible tener una lista de todas las viviendas, pero obtener una lista de las personas que las habitan sería muy difícil. Por lo tanto, el muestreo por conglomerados es ampliamente aplicado en estudios de muestreo.

El muestreo estratificado se diferencia del muestreo por conglomerados en que del primero se extrae una muestra a **todos los estratos** y en el segundo, después de dividir la población, se extrae una **muestra de subpoblaciones o conglomerados**; posteriormente, según sea el caso, se censa (se toma completo el conglomerado) o se extrae una muestra.

No se debe olvidar que la condición necesaria para que el muestreo por conglomerados sea válido, es que todos los elementos que constituyen la población deben de pertenecer a uno y sólo a un conglomerado y, además, la unión de todos vuelve a conformar todas las unidades de muestreo de la población; esto evita que se cometan errores de omisión y duplicación. Si la condición no se llegase a cumplir entonces habrá sesgos. No es necesario que todos los elementos asociados con un conglomerado tengan que estar localizados físicamente dentro de sus fronteras. A este tipo de conglomerados se les conoce como conglomerados abiertos.

Otra observación es que el tamaño del conglomerado influye en el muestreo. En general, entre más pequeña sea el área, más precisa será la estimación de la población que se esté encuestando.

Por ejemplo, una selección de viviendas aleatoria e independiente, probablemente estará más dispersa sobre una gran ciudad (lo que provocaría más cruces de secciones en la población), que una muestra con el mismo número de predios pero conglomeradas, en donde para cada conglomerado tendríamos varias viviendas juntas en pequeños barrios o vecindarios. Por otro lado sería más costosa una muestra ampliamente dispersa sobre la población total de edificios y casas, que una muestra equivalente por conglomerados, puesto que el costo adicional de encuestar dos edificios vecinos es menor al de encuestar a los dos edificios por separado y además considerarlos como independientes.

El conglomerado óptimo es aquel que estima la característica o atributo a estudiar con la mayor precisión para una proporción de la población dada o en términos más reales, para un costo dado.

2.2.2 Notación

Suponiendo que la población está dividida en N conglomerados constituidos por M elementos; además que seleccionamos n conglomerados de los N por MAS. Se denota a

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

$$\bar{Y}_{i\cdot} = \frac{1}{M} \sum_{j=1}^M Y_{ij}, \text{ la media particular de cualquier conglomerado} \quad (2.13)$$

$$\bar{\bar{Y}} = \frac{1}{N} \sum_{i=1}^N \bar{Y}_{i\cdot}, \text{ la media de las medias particulares que se obtuvieron de cada conglomerado} \quad (2.14)$$

$$\bar{Y} = \frac{1}{NM} \sum_{i=1}^N \sum_{j=1}^M Y_{ij}, \text{ la media poblacional} \quad (2.15)$$

$$S_i^2 = \frac{1}{(M-1)} \sum_{j=1}^M (Y_{ij} - \bar{Y}_{i\cdot})^2, \text{ cuadrados medios del } i\text{-ésimo conglomerado} \quad (2.16)$$

$$S_w^2 = \frac{1}{N} \sum_{i=1}^N S_i^2, \text{ varianza dentro de los conglomerados} \quad (2.17)$$

$$S_b^2 = \frac{1}{N-1} \sum_{i=1}^N (\bar{Y}_{i\cdot} - \bar{\bar{Y}})^2, \text{ varianza entre conglomerados} \quad (2.18)$$

$$S^2 = \frac{1}{NM-1} \sum_{i=1}^N \sum_{j=1}^M (Y_{ij} - \bar{Y})^2, \text{ varianza poblacional} \quad (2.19)$$

ESTIMACIONES

Un estimador de \bar{Y} es la media muestral de las medias muestrales particulares \bar{y} expresado

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n \bar{y}_{i\cdot} \quad (2.20)$$

la varianza del estimador anterior es

$$V(\bar{y}) = \left(\frac{1}{n} - \frac{1}{N} \right) S_b^2 \quad (2.21)$$

como s_b^2 es un estimador de S_b^2 , entonces un estimador de $V(\bar{y})$ es

$$\hat{V}(\bar{y}) = \left(\frac{1}{n} - \frac{1}{N}\right)s_b^2 \quad (2.22)$$

tanto \bar{y} como $\hat{V}(\bar{y})$ son estimadores sin sesgo (Sukhatme et al, 1984 pág 272).

2.3 Submuestreo o Muestreo Bietápico

2.3.1 Descripción

Hasta aquí se han considerado solamente técnicas de muestreo en las que el conglomerado seleccionado ha sido censado; asimismo, que entre mayor sea el tamaño del conglomerado, la eficiencia del proceso decrece, siempre y cuando los elementos sean las unidades de muestreo. Así, es lógico esperar que para un número dado de elementos, se obtendrá mayor precisión al distribuirlos en varios conglomerados, que tomar un pequeño número de éstos o extraer muestras muy grandes de cada uno.

Al método de escoger primero conglomerados y luego seleccionar un número específico de cada uno se le conoce como submuestreo, también llamado muestreo en dos etapas o bietápico. Los conglomerados que constituyen las unidades de muestreo en la primera etapa reciben la denominación de unidades de primera etapa o primarias; por su parte, los elementos o grupos de elementos que se encuentran dentro de cada conglomerado, constituyen las unidades de muestreo de la segunda etapa, también conocidas como subunidades o unidades secundarias. Este método puede realizarse en tres o más etapas, en cuyo caso recibe el nombre de muestreo trietápico o multietápico respectivamente.

Un ejemplo de muestreo trietápico, lo proporciona un estudio sobre el rendimiento de las cosechas en una determinada región, en el que las aldeas representan las unidades primarias, los terrenos de siembra que se encuentran dentro de las aldeas constituyen unidades de la segunda etapa y, las parcelas o pequeñas extensiones de tierra que componen cada terreno, las unidades de la tercera etapa. Además, sería innecesario y poco rentable contabilizar la producción total del terreno, porque los rendimientos de parcelas contiguas dentro de un mismo campo de cultivo son semejantes.

2.3.2 Notación

Es posible que las unidades primarias sean todas del mismo tamaño o bien, que sean desiguales, cada caso tiene su propia notación, en este caso utilizaremos unidades de distintos tamaños.

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

$$\bar{M} = \frac{1}{N} \sum_{i=1}^N M_i, \text{ tamaño promedio de unidades de segunda etapa} \quad (2.23)$$

$$\bar{Y}_{i\cdot} = \frac{1}{M_i} \sum_{j=1}^{M_i} Y_{ij}, \text{ la media de las observaciones en la } i\text{-ésima unidad primaria} \quad (2.24)$$

$$\bar{Y} = \bar{Y}_{\cdot\cdot} = \frac{1}{NM} \sum_{i=1}^N M_i \bar{Y}_{i\cdot}, \text{ la media poblacional} \quad (2.25)$$

$$S_{iy}^2 = \frac{1}{(M_i - 1)} \sum_{j=1}^{M_i} (Y_{ij} - \bar{Y}_{i\cdot})^2 \quad (2.26)$$

$$S_{by}^2 = \frac{1}{(N - 1)} \sum_{i=1}^N \left(\frac{M_i}{M} \bar{Y}_{i\cdot} - \bar{Y} \right)^2 \quad (2.27)$$

$$U_{ij} = \frac{M_i}{M} Y_{ij} \quad (2.28)$$

$$\bar{U}_{i\cdot} = \frac{1}{M_i} \sum_{j=1}^{M_i} U_{ij} = \frac{M_i}{M} \bar{Y}_{i\cdot} \quad (2.29)$$

$$\bar{U} = \frac{1}{N} \sum_{i=1}^N \bar{U}_{i\cdot} = \bar{Y}_{\cdot\cdot} \quad (2.30)$$

$$S_{iu}^2 = \frac{1}{(M_i - 1)} \sum_{j=1}^{M_i} (U_{ij} - \bar{U}_{i\cdot})^2 = \left(\frac{M_i}{M} \right)^2 S_{iy}^2 \quad (2.31)$$

$$S_{iu}^2 = \frac{1}{N - 1} \sum_{i=1}^N (\bar{U}_{i\cdot} - \bar{U})^2 = S_{by}^2 \quad (2.32)$$

m_i = el número de unidades de la segunda etapa elegidas en la i -ésima unidad primaria muestreada

$$\bar{m} = \frac{1}{n} \sum_{i=1}^n m_i \quad (2.33)$$

$$\bar{y}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} y_{ij} \quad (2.34)$$

$$\bar{u}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} u_{ij} = \frac{1}{m_i} \sum_{j=1}^{m_i} \frac{M_i}{M} Y_{ij} \quad (2.35)$$

$$\bar{u} = \frac{1}{n} \sum_{i=1}^n \bar{u}_i = \frac{1}{n} \sum_{i=1}^n \frac{M_i}{M} \bar{y}_i \quad (2.36)$$

ESTIMACIONES

Existen varios estimadores de la media poblacional para este diseño de muestreo, aquí se mencionan tres de los cuatro que presenta Sukhatme et al (1984, pp. 323-336).

(i) \hat{Y}_2 es un estimador sesgado de \bar{Y}

$$\hat{Y}_2 = \frac{1}{n} \sum_{i=1}^n \bar{y}_i \quad (2.37)$$

$$V(\hat{Y}_2) = \left(\frac{1}{n} - \frac{1}{N}\right) \frac{1}{(N-1)} \sum_{i=1}^N \left(\bar{Y}_i - \frac{1}{N} \sum_{i=1}^N \bar{Y}_i\right)^2 + \frac{1}{nN} \sum_{i=1}^N \left(\frac{1}{m_i} - \frac{1}{M_i}\right) S_{iy}^2 \quad (2.38)$$

donde S_{iy}^2 es (2.26)

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

$$\hat{Y}_{u2} = \frac{1}{n} \sum_{i=1}^n \frac{M_i}{M} \bar{y}_i = \frac{1}{n} \sum_{i=1}^n \bar{u}_i = \bar{u} \quad (2.39)$$

$$V(\hat{Y}_{u2}) = \left(\frac{1}{n} - \frac{1}{N}\right) \frac{1}{(N-1)} \sum_{i=1}^N \left[\left(\frac{M_i}{M}\right) \bar{Y}_i - \bar{Y}\right]^2 + \frac{1}{nN} \sum_{i=1}^N \left(\frac{1}{m_i} - \frac{1}{M_i}\right) \left(\frac{M_i}{M}\right)^2 S_{iy}^2 \quad (2.40)$$

el estimador de la varianza es

$$\hat{V}(\hat{Y}_{u2}) = \left(\frac{1}{n} - \frac{1}{N}\right) s_u^2 + \frac{1}{nN} \sum_{i=1}^n \left(\frac{1}{m_i} - \frac{1}{M_i}\right) s_{iu}^2 \quad (2.41)$$

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

$$\hat{Y}_{r2} = \frac{\sum_{i=1}^n M_i \bar{y}_i}{\sum_{i=1}^n M_i} \quad (2.42)$$

la varianza aproximada del estimador está dada por

$$V_1\left(\hat{Y}_{r2}\right) = \frac{N-n}{Nn} S_{by}^2 \quad (2.43)$$

donde S_{by}^2 es (2.27), el estimador de la varianza es

$$\hat{V}_1 = \left(\hat{Y}_{r2}\right) = \frac{N-n}{Nn} S_{by}^2 \quad (2.44)$$

donde

$$s_{by}^2 = \frac{1}{n-1} \sum_{i=1}^n \frac{M_i^2}{M_n^2} (\bar{y}_i - \bar{u})^2 \quad (2.45)$$

donde \bar{u} es (2.36) y

$$\bar{M}_n = \frac{1}{n} \sum_{i=1}^n M_i \quad (2.46)$$

El sesgo del primer estimador es provocado porque las probabilidades de selección de las subunidades varían de una unidad primaria a otra, esto se debe precisamente a que las unidades primarias son de distinto tamaño. Ahora bien, si M_i tiene una variación considerable y además está correlacionada con la característica bajo estudio, el sesgo se puede despreciar pero, es muy importante asegurarse de que estas condiciones se cumplan antes de aceptar que \hat{Y}_2 es un estimador de \bar{Y} .

Hablando del segundo estimador, el primer sumando de (2.41) depende únicamente de la variación de los totales entre conglomerados. Comparando con el primer estimador se puede demostrar que el primer sumando de (2.41) es mayor casi siempre que el primer sumando del estimador sesgado es decir (2.38), con las condiciones de que la correlación entre el tamaño del conglomerado y su media sea positiva, y que el sesgo $(E(\hat{Y}_2) - \bar{Y})$ sea insignificante. De igual forma el segundo sumando de (2.41) será mayor al de (2.38) a condición de que exista correlación positiva entre M_i y S_i^2 .

El tercer estimador es una media ponderada de medias de conglomerados; está fundamentado en métodos de estimadores de razón porque se compone de la razón de dos variables aleatorias. Este estimador también es sesgado pero el sesgo se reduce conforme n crece; otra ventaja es que no es necesario conocer \bar{M} para calcularlo.

"Por lo general, cuando el tamaño de los conglomerados (M_i) varía considerablemente, quizá el estimador \hat{Y}_{r2} es el más eficiente de los cuatro, siempre y cuando: a) n sea grande y b) x^3 y y tengan un alto grado de correlación". (Sukhatme 1984, pp. 333).

El otro estimador no mencionado aquí es el de regresión. Ver Sukhatme et al (1984, pp. 290-293 y 333-336).

2.4 Submuestreo Estratificado

2.4.1 Descripción

En el muestreo bietápico es posible aplicar otro tipo de técnicas sobre las unidades de la primera etapa, y es muy común encontrarse con diseños que combinan diversos procesos; en este caso tenemos relacionados el muestreo de dos etapas y el muestreo estratificado, habiendo sido descritos anteriormente por separado, sólo resta explicar la forma en que interactúan: se estratifica a la población; después, dentro de cada estrato se realiza un muestreo bietápico, tomando una muestra aleatoria de unidades de primera etapa. Para obtener las unidades secundarias se extrae otra muestra aleatoria de elementos de cada unidad primaria seleccionada. Para mayor claridad ver los estratos, las unidades primarias y secundarias de la encuesta preelectoral en la sección 3.1.3.

2.4.2 Notación

Supongamos una población dividida en L estratos:

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

$$N = \sum_{t=1}^L N_t \quad (2.47)$$

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

$M_{it} = N_t \bar{M}_t = \sum_{j=1}^{M_{it}} M_{ij}$ = número total de unidades secundarias en el t -ésimo estrato

Y_{ij} = valor del atributo en la j -ésima unidad secundaria de la i -ésima unidad primaria dentro del t -ésimo estrato

$$\bar{Y}_{it} = \frac{1}{M_{it}} \sum_{j=1}^{M_{it}} Y_{ij} \quad (2.48)$$

$$\bar{Y}_{it} = \frac{1}{M_{it}} \sum_{j=1}^{M_{it}} Y_{ij} = \frac{1}{N_t \bar{M}_t} \sum_{i=1}^{N_t} \sum_{j=1}^{M_{it}} Y_{ij} \quad (2.49)$$

$$\bar{Y} = \bar{Y}_{\dots} = \frac{\sum_{t=1}^L M_{t0} \bar{Y}_{it}}{\sum_{t=1}^L M_{t0}} = \sum_{t=1}^L \lambda_t \bar{Y}_{it} \quad (2.50)$$

$$\lambda_t = \frac{M_{t0}}{\sum_{t=1}^L M_{t0}} = \frac{M_{t0}}{M_0} \quad (2.51)$$

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

$$n = \sum_{t=1}^L n_t \quad (2.52)$$

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

$$S_{iy}^2 = \frac{1}{(M_{ii} - 1)} \sum_{i=1}^{N_t} (Y_{ij} - \bar{Y}_{i..})^2 \quad (2.53)$$

$$S_{ib}^2 = \frac{1}{(N_t - 1)} \sum_{i=1}^{N_t} \left(\frac{M_{ii}}{M_t} \bar{Y}_{i..} - \bar{Y}_{t..} \right)^2 \quad (2.54)$$

$$\bar{y}_{ii} = \frac{1}{m_{ii}} \sum_{j=1}^{m_{ii}} y_{ij} \quad (2.55)$$

$$\bar{y}_{it} = \frac{1}{M_{it} N_t} \sum_{i=1}^{m_{ii}} \sum_{j=1}^{m_{ij}} y_{ij} = \frac{1}{N_t} \sum_{i=1}^{m_{ii}} \frac{M_{ii}}{M_t} \bar{y}_{ii} \quad (2.56)$$

$$\bar{y}_{w2} = \sum_{i=1}^L \lambda_i \bar{y}_{it} \quad (2.57)$$

$$s_{iy}^2 = \frac{1}{(m_{ii} - 1)} \sum_{j=1}^{m_{ii}} (v_{ij} - \bar{v}_{i.})^2 \quad (2.58)$$

$$s_{ib}^2 = \frac{1}{(n_t - 1)} \sum_{i=1}^{n_t} \left(\frac{M_{ii}}{M_t} \bar{y}_{i.} - \bar{y}_{t.} \right)^2 \quad (2.59)$$

ESTIMACIONES

La expresión (2.57) es un estimador de \bar{Y} con la siguiente varianza

$$V(\bar{y}_{w2}) = \sum_{i=1}^L \lambda_i^2 \left[\left(\frac{1}{m_{ii}} - \frac{1}{N_t} \right) S_{iy}^2 + \frac{1}{n_t N_t} \sum_{i=1}^{n_t} \left(\frac{M_{ii}}{M_t} \right)^2 \left(-\frac{1}{m_{ii}} - \frac{1}{M_t} \right) S_{ib}^2 \right] \quad (2.60)$$

el estimador de la varianza anterior es

$$\hat{v}(\bar{y}_{w2}) = \sum_{i=1}^L \lambda_i^2 \left[\left(\frac{1}{m_{ii}} - \frac{1}{N_t} \right) s_{iy}^2 + \frac{1}{n_t N_t} \sum_{i=1}^{n_t} \left(\frac{1}{m_{ii}} - \frac{1}{M_t} \right) \left(\frac{M_{ii}}{M_t} \right)^2 s_{ib}^2 \right] \quad (2.61)$$

ambos estimadores son insesgados. Sukhatme et al, (1984 pp. 353-354).

La expresión (2.60) se puede escribir en forma resumida utilizando (2.40) en cada estrato

$$V(\bar{y}_{w2}) = \sum_{i=1}^L \lambda_i^2 V(\bar{y}_{it}) \quad (2.62)$$

lo mismo sucede con (2.61) pero, utilizando (2.41) y $\hat{v}(\bar{y}_{it})$

$$\hat{v}(\bar{y}_{w2}) = \sum_{i=1}^L \lambda_i^2 \hat{v}(\bar{y}_{it}) \quad (2.63)$$

$$\text{donde} \quad \hat{V}(\bar{y}_{st}) = \left(\frac{1}{n_t} - \frac{1}{N_t}\right) s_{yt}^2 + \frac{1}{n_t N_t} \sum_{u=1}^{U_t} \left(\frac{1}{m_u} - \frac{1}{M_u}\right) \left(\frac{M_u}{M_t}\right)^2 s_{uy}^2 \quad (2.64)$$

2.5 Muestreo por Cuotas

2.5.1 Descripción

Debido a la carencia de literatura en español de una descripción del muestreo por cuotas, a continuación se presenta la traducción sobre el tema de Stephan et al (1958, capítulo 3, pp. 37-38)⁴. Sin embargo, si existe una conceptualización muy clara y sencilla sobre este método en: Castrejón J. L.⁵ (Tres Encuestas por Muestreo (Tesina), 1996, pp. 25-26). Al final de la misma se hará un comentario basado en la traducción y en la obra citada.

Desde hace mucho tiempo una variante del diseño de muestreo que se usa con mucha frecuencia ha sido el muestreo por cuota. Fue desarrollado por Cherington, Roper, Gallup y Crossley cuyos estudios de opinión fueron difundidos muy ampliamente después de la elección presidencial de Estados Unidos de 1936, además de otros realizados por bufetes de investigaciones de mercados que no gozaron de la misma difusión.

Poseía características tan simples que de inmediato fue aceptado; por un lado, porque incluirlo en una encuesta era como estar a la moda, ya que otras encuestas famosas lo habían utilizado, y por el otro, ofrece grandes atractivos con respecto a costos y conveniencia. En años más recientes ha sido reemplazado por otras técnicas de "muestreo probabilístico".

Como la mayoría de los diseños de muestreo, el muestreo por cuota, parte de la premisa de que la muestra debe estar dispersa sobre la población y que debe contener la misma fracción de individuos que poseen ciertas características propias de la población. La información censal y de otras encuestas son utilizadas para dividir a la población en cierto número de subpoblaciones mutuamente exclusivas o

⁴ Stephan F. y McCarthy P., *Sampling Opinions*, 1958 New York John Wiley & Sons
⁵ Castrejón Caballero J. L., *Tres Encuestas por Muestreo* (Tesina), 1996 Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, U.N.A.M.

estratos y también para la determinación o estimación (dependiendo de para qué estén adecuados los datos) de la fracción de población en cada uno de los estratos.

La muestra es asignada a estratos tales como regiones geográficas, grupos por edad y sexo, grupos raciales, etc.; esta asignación va en proporción al tamaño real o estimado de cada estrato.

La aplicación de controles o filtros sobre la muestra, que en ocasiones son externos a la encuesta, llevan a cabo implícitamente una estratificación, incluso en otros tipos de diseños muestrales. Por lo tanto, es muy importante aclarar que, decir que el muestreo por cuota es un muestreo estratificado, es arriesgado, porque no es lo suficientemente poderoso para distinguirlo de otros procedimientos de muestreo.

Al respecto existen muchas observaciones que resaltar, como por ejemplo, ¿cuántos estratos serán usados? ¿cómo se definirá un estrato? ¿cómo será asignada la muestra en cada uno? ¿cómo se practicará la muestra dentro de cada estrato?

El paso final en el muestreo por cuota es la distribución de la muestra entre los enumeradores o encuestadores, y asignarles a cada uno el número de personas que deben encuestar en cada uno de los estratos. Así sabrán a cuántos hombres y mujeres o cuántas 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 contáramos con la plena 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, entonces el muestreo por cuota podría llamarse muestreo aleatorio estratificado. En la actualidad, es sumamente difícil saber si los enumeradores llenaron su cuota de acuerdo a los esquemas del plan para la selección.

Hay muchos factores que influyen al enumerador para seleccionar a sus respondientes, tales como preferencias conscientes o inconscientes, gustos, hábitos en fin. Por ejemplo, el entrevistador puede sentirse atraído o rechazado por la simple apariencia de una casa; o tal vez puede concentrar más su trabajo en ciertas partes de la ciudad y descuidar otras; según su criterio puede omitir a cierto tipo de personas porque puede considerar que no son convenientes para la entrevista.

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. Se han hecho varias modificaciones al método, las cuales están enfocadas directamente hacia la intención de que el encuestador se convierta en una máquina en el momento de hacer la selección y olvide sus preferencias. Esta variante quizá haga más específica y definida la selección o por lo menos, hará que se reduzca al mínimo el efecto de subjetividad que existe cuando se realiza la entrevista. El entrenamiento que se les imparte a los enumeradores tiene varios objetivos, entre otros tenemos el de no omitir algunos grupos, que pueden ser los marginados y los de altos ingresos; el de cómo deben hacer las preguntas de la encuesta, si se trata de una casa habitación o de una granja; y tal vez el más valioso, es el hacer consciencia de la necesidad de lograr llenar un cuestionario con un mínimo de irregularidades.

Es posible reducir los grupos en los que se aplicaran las cuotas, esto hará que sean más específicos. Por ejemplo, señalándole a los entrevistadores los lugares donde deben realizar la encuesta tales como: una manzana específica de una colonia o barrio específico, en secciones claramente definidas dentro de la ciudad.

Existen distintos modos de interpretar las consideraciones anteriores para definir el proceso del muestreo por cuota, es más, incluso hay otros que no fueron tratados aquí. Por lo tanto, no es suficiente que con el simple hecho de manifestar que se utilizó el muestreo por cuota en una encuesta, esperemos que cualquier persona tenga más claro el hecho de cómo se efectúa la extracción de una muestra.

Entonces, antes de afirmar que por lo menos dos de estos procedimientos son burdamente parecidos, debemos describir con gran claridad y en detalle el desarrollo del método.

Ciertamente el muestreo por cuotas era novedoso en la década de los cincuentas, tal vez esto se debía a los costos bajos con los que operaba, y éste es un aspecto fundamental en el uso del muestreo. Por otro lado, es extraño que los autores sitúen entre comillas al muestreo probabilístico, porque en la actualidad se maneja como un término natural entre los muestristas.

De cualquier manera este es el punto coyuntural del muestreo por cuotas, porque el espíritu de seleccionar las muestras de manera aleatoria desaparece, por lo que la mayor parte de la selección queda en manos del encuestador; se menciona la mayor parte, porque algo debe recaer en los supervisores y en las personas que instruyen a los encuestadores. Entonces se presenta el efecto de la subjetividad que influye en cada enumerador.

Es posible perfeccionar el muestreo por cuotas, pero, el proceso de perfeccionamiento seguramente encarecerá el diseño. Como toda la selección es a criterio del enumerador, una solución sería la profesionalización de los entrevistadores; es decir, que no se contrate a cualquier persona para realizar la encuesta, sino contratar o contar con un grupo propio de encuestadores previamente conscientizados y altamente calificados para hacer el trabajo; es obvio que al tener personal profesional se esfuma la idea de los costos bajos del muestreo por cuotas pero, es muy probable que los sesgos provocados por el efecto subjetivo se reduzcan y, al mismo tiempo, se incremente la eficiencia y precisión de las estimaciones.

CAPITULO 3 Encuesta Preelectoral

3.1 Introducción

En este capítulo se analizará la encuesta preelectoral del 14 de mayo de 1995, de las elecciones de carácter extraordinario para elegir gobernador de Guanajuato convocadas para el 28 de mayo de 1995; a través de un enfoque propiamente descriptivo. El capítulo está dividido en dos apartados; el primero, llamado introducción, contiene cuatro secciones donde son descritas la encuesta preelectoral y la logística del proceso, como son el diseño muestral, la estratificación y el operativo de campo. El segundo, contiene el análisis de los resultados, esto lo hace muy extenso porque aparte de la información electoral crea su propia información sociodemográfica. Se compone de siete secciones de las cuales una tiene dos subsecciones claramente identificadas.

La encuesta preelectoral que se analizará es la segunda de las dos que llevó a cabo el CEO; de la primera (levantada el 20 de abril de 1995) no se realizará ningún análisis.

3.1.1 Descripción

Una encuesta preelectoral es un indicador de la popularidad o fuerza de simpatía que poseen un candidato y un partido, pero por complemento también indica el polo opuesto, es decir, la impopularidad o fuerza de repulsión hacia el partido o el individuo.

Al llevar a cabo encuestas de este tipo se mide la popularidad y la impopularidad de una persona que es candidata a algún puesto específico, aunque realmente lo importante es saber que proporción de la población la apoya, y también es interesante saber el porqué.

Una encuesta preelectoral posee tres preguntas fundamentales que son: ¿por cuál contendiente van a votar?, ¿conocen su desarrollo político? y ¿por qué votan por él?. Pero no es sólo el individuo, sino, más bien, es una agrupación la que está postulando a dicho individuo para desempeñar el puesto, esta agrupación es un partido político. Como una identidad, un partido político también es susceptible de ser medido en cuanto a su popularidad o lo contrario, y a final de cuentas ésta es la información que interesa dentro de una encuesta preelectoral, y no tanto la del propio candidato. Una persona puede ser muy bien aceptada a nivel masivo por su propia simpatía, pero, en el momento en que se relaciona con una agrupación que es impopular dentro del mismo contexto, la fuerza de simpatía propia empieza a ser cuestionada, incluso por aquellos que lo aprobaban.

Hasta este momento se ha hablado sólo de un candidato y de un partido político, pero, ¿qué sucede, si existen más candidatos que pertenezcan a otras facciones, o candidatos independientes que también gocen a nivel poblacional de una atracción significativa?.

Es entonces, cuando una encuesta preelectoral entra en acción, porque indicará lo popular que es cada candidato que contiene en un proceso electoral. Pero única y exclusivamente en ese momento, y aunque esta medición se realiza muy cerca del periodo de elecciones, por ningún motivo puede calificarse o utilizarse como un predicción de resultados finales.

Después de esta breve descripción, respecto a los candidatos cabe mencionar que la encuesta también investiga sobre otras variables que están relacionadas con el proceso electoral, y otras que tienen mínima relación, pero son significativas para el aspecto social que prevalece antes de las elecciones. Más adelante, dentro de este capítulo se exponen estas variables.

3.1.2 Diseño Muestral

En la presente encuesta preelectoral, se aplicó la técnica del muestreo estratificado de dos etapas o bietápico; las razones por las que se decidió tal procedimiento fueron la operatividad, el costo y la representatividad en cuanto a opinión se refiere, porque el proceso requiere que los estimadores sean lo más preciso posible y la naturaleza propia de la investigación así lo exige; se deben aprovechar los recursos con los que se cuenta, tales como la información que proporciona el INEGI, de donde se obtuvo el número, el nombre y la población mayor de 18 años de edad de cada municipio; con esta información se logra la estratificación de la población a estudiar.

Guanajuato esta conformado por 46 municipios oficiales, los cuales fueron divididos en 7 estratos. El estrato número uno lo integran 12 municipios cuya población es menor a diez mil habitantes (10,000); el estrato número dos comprende 16 municipios los cuales tienen una población entre los diez mil habitantes y los treinta mil (de 10,001 a 30,000); el número tres tiene 14 municipios cuyo tamaño está entre treinta mil uno y setenta mil habitantes (de 30,001 a 70,000); el resto de los estratos son los cuatro municipios más grandes del estado cuyas principales ciudades llevan el mismo nombre; con la siguiente asignación numérica: cuatro para Salamanca, cinco para Celaya, seis para Irapuato y siete para León.

Una vez definidos los estratos, en cada uno se aplicó el muestreo bietápico, donde las unidades primarias serán los municipios; las secundarias, los habitantes mayores de 18 años residentes en él. La distribución de la muestra entre los estratos se prorrateó de manera proporcional a su tamaño. Del estrato uno (12 municipios) se extrajo una muestra aleatoria de 6 municipios; dentro del segundo (16 municipios), se tomó una muestra de 12 y en los estratos restantes se consideró su totalidad. La tabla 1 de la página 41 muestra la distribución.

La muestra asignada a cada estrato se distribuyó proporcionalmente al tamaño del municipio muestreado. La selección de los individuos dentro de cada municipio se hizo bajo el criterio de cuotas; los criterios de cuota fueron sexo y grupo de edad (18-24, 25-34, 35-44, 45 y más) considerados de manera simultánea y en porcentajes similares a los de la distribución poblacional.

3.1.3 Aplicación del Muestreo Estratificado Bietápico

Las unidades primarias serán los municipios y las secundarias los ciudadanos residentes en cada municipio que, desde este momento llamaremos respondientes; no se debe olvidar que de cada estrato se puede extraer una muestra, o bien, es posible tomar a todas las unidades que lo integran. En este caso, todos los municipios del estrato número tres fueron seleccionados. Haciendo la adecuación de la notación de la sección 2.4.2 del capítulo 2 para este diseño muestral en particular, se presenta la notación siguiente.

M = Número de habitantes en la población total, esto es, solo la población guanajuatense con derecho a voto

N = Número de municipios

N_t = Número de municipios en el t -ésimo estrato de la población, $t=1, \dots, k$

n = Número de municipios en muestra

n_t = Número de municipios muestreados en el estrato t , $t=1, \dots, k$

k = Número de estratos

M_{it} =Número de habitantes en el municipio i del estrato t , $i=1, \dots, N$; $t=1, \dots, k$

m_{it} =Número de habitantes muestreados en el municipio i del estrato t , $i=1, \dots, N$; $t=1, \dots, k$

\bar{M}_t = Tamaño promedio de municipio en el estrato t , $t=1, \dots, k$

M_{t0} = Número de respondientes en el estrato t , $t=1, \dots, k$

Con t se representa el índice que recorrerá los estratos en cuestión, pero como $k=7$, entonces t solo tomará los valores de 1 a 7, consecuentemente; $t=1, \dots, 7$.

El índice que recorrerá a los municipios será i , por lo que i tomará los valores desde 1 hasta N , de igual forma $N=46$.

También se deben readaptar las ecuaciones dadas en la sección 2.4.1, las cuales se muestran en seguida. [se debe recordar que la notación es la adoptada de Sukhatme, (1984, pp. 352 - 353)].

$$N = \sum_{t=1}^k N_t \quad (3.1)$$

$$n = \sum_{t=1}^k n_t \quad (3.2)$$

$$\bar{M}_t = \frac{\sum_{i=1}^{N_t} M_{it}}{N_t} \quad (3.3)$$

$$M_{t0} = \sum_{i=1}^{N_t} M_{it} \quad (3.4)$$

$$M = \sum_{t=1}^k M_{t0} \quad (3.5)$$

El valor medio muestral de la característica de interés (voto válido) en el estrato t está expresado como sigue:

$$\bar{y}_t = \frac{1}{n_t} \sum_{i=1}^{n_t} \frac{M_{it}}{M_t} \frac{1}{m_{it}} \sum_{j=1}^{m_{it}} y_{tij} \quad (3.6)$$

El estimador de la media poblacional es la suma ponderada de las medias por municipio, expresado así

$$\bar{y}_w = \sum_{t=1}^k \frac{M_{it}}{M} \bar{y}_t \quad (3.7)$$

Si sustituimos (3.6) en (3.7) y hacemos las operaciones algebraicas necesarias, el estimador poblacional se puede expresar como una suma ponderada de las medidas originales, esta suma es:

$$y_w = \sum_{t=1}^k \sum_{i=1}^{n_t} \sum_{j=1}^{m_{it}} \frac{N_t M_{it}}{n_t m_{it} M} y_{tij} \quad (3.8)$$

donde

$$y_{tij} = \begin{cases} 1 & \text{si el entrevistado contesta afirmativamente la pregunta en cuestión} \\ 0 & \text{si el entrevistado contesta negativamente} \end{cases}$$

$t = 1, \dots, k; i = 1, \dots, n_t; j = 1, \dots, m_{it}$

La varianza del estimador esta dada por la siguiente expresión:

$$\text{var}(\bar{y}_w) = \sum_{t=1}^k \lambda_t^2 \left[\left(\frac{1}{n_t} - \frac{1}{N_t} \right) S_{it}^2 + \frac{1}{n_t N_t} \sum_{i=1}^{N_t} \frac{M_{it}^2}{M_t^2} \left(\frac{1}{m_{it}} - \frac{1}{M_{it}} \right) S_{it}^2 \right] \quad (3.9)$$

donde

$$S_{it}^2 = \frac{1}{N_t - 1} \sum_{i=1}^{N_t} \left(\frac{M_{it}}{M_t} \bar{y}_{i(M_t)} - \bar{y}_{i(M_t)} \right)^2 \quad (3.10)$$

$$S_{it}^2 = \frac{1}{M_{it} - 1} \sum_{j=1}^{M_{it}} \left(y_{tij} - \bar{y}_{i(M_t)} \right)^2 \quad (3.11)$$

$$\lambda_t = \frac{M_{it}}{M} \quad (3.12)$$

$$\bar{y}_{i(M_t)} = \frac{1}{M_{it}} \sum_{j=1}^{M_{it}} y_{tij} \quad (3.13)$$

$$\bar{y}_{i(M_t)} = \frac{1}{M} \sum_{i=1}^{N_t} \sum_{j=1}^{M_{it}} y_{tij} \quad (3.14)$$

Los estimadores para S_{ih}^2 y S_{ii}^2 son respectivamente

$$\hat{S}_{ih}^2 = \frac{1}{n_i - 1} \sum_{i=1}^{n_i} \left(\frac{M_{ii}}{M_i} \bar{y}_{ii(m_i)} - \bar{y}_{iim} \right)^2 \quad (3.15)$$

$$\hat{S}_{ii}^2 = \frac{1}{m_{ii} - 1} \sum_{j=1}^{m_{ii}} \left(y_{ij} - \bar{y}_{ii(m_i)} \right)^2 \quad (3.16)$$

donde

$$\bar{y}_{ii(m_i)} = \frac{1}{m_{ii}} \sum_{j=1}^{m_{ii}} y_{ij} \quad (3.17)$$

El estimador de la varianza lo obtenemos al operar la suma que está dentro de los paréntesis cuadrados de la expresión (3.9) hasta n_i , en vez de operarla hasta N_i , y también al sustituir (3.15) y (3.16) por S_{ih}^2 y S_{ii}^2 respectivamente en la misma, por lo que (3.9), después de las modificaciones citadas, le da forma a la nueva expresión (Sukhatme, 1984, pp. 352- 354)

$$\hat{v}ar(\bar{y}_w) = \sum_{i=1}^k \lambda_i^2 \left[\left(\frac{1}{n_i} - \frac{1}{N_i} \right) \hat{S}_{ih}^2 + \frac{1}{n_i N_i} \sum_{i=1}^{n_i} \frac{M_{ii}^2}{M_i^2} \left(\frac{1}{m_{ii}} - \frac{1}{M_{ii}} \right) \hat{S}_{ii}^2 \right] \quad (3.18)$$

Por último, se construyen los intervalos para la proporción poblacional de cierta característica, para aquellas proporciones entre 0.3 y 0.7 expresado así

$$\left(\bar{y}_w - t \sqrt{\hat{v}ar(\bar{y}_w)}, \bar{y}_w + t \sqrt{\hat{v}ar(\bar{y}_w)} \right) \quad (3.19)$$

donde t es el desvío normal correspondiente al nivel de confianza

3.1.4 Operativo de Campo

El tamaño de la muestra se calculó considerando dos aspectos:

- a) Bajo un esquema de muestreo aleatorio simple, se tiene un error relativo de 5% a un nivel de confianza de 95%.
- b) Los resultados se puedan comparar con una primera encuesta preelectoral aplicada el 20 de abril de 1995 (no presentada aquí y bajo otro diseño muestral), en cuanto a tamaño de muestra.

La distribución de la muestra pretende asegurar los puntos siguientes:

- a) Representatividad de municipios de diversos tamaños de población, (pequeños, medianos, grandes, y de los cuatro municipios que contienen a las cuatro ciudades mayores: Salamanca, Celaya, Irapuato y León).
- b) Esparcimiento geográfico-poblacional.

Operativamente las encuestas por municipio se aplicaron así:

- a) Los municipios grandes se dividieron en cuadrantes y se fijaron 5 áreas de muestreo, una para cada cuadrante y otra en el centro del municipio. Dentro de cada área se fijaron varios puntos muestrales.
- b) En municipios pequeños las encuestas se aplicaron básicamente en el centro del municipio.

Se decidió que el tamaño de muestra se basara en un diseño de MAS por las siguientes explicaciones:

- I) Se desconocen los estimadores para las varianzas de los estimadores en los municipios y para cada estrato.
- II) El muestreo estratificado con asignación proporcional es más preciso que el MAS, sin embargo, el muestreo bietápico es menos preciso con respecto al MAS.

Supongamos que p denota el estimador de la proporción P de individuos pertenecientes a una clase C . Vamos a calcular el tamaño de muestra requerido para estimar P con un grado de confiabilidad α y un error absoluto d .

Cochran (1990, pág. 107), nos lo explica muy sencillamente: " Se ha convenido en algún margen de error d de la proporción estimada p de las unidades en la clase C , y existe un pequeño riesgo α , que estamos dispuestos a correr, de que el error real supere a d ; es decir, queremos que"

$$\Pr(|p - P| \geq d) = \alpha$$

Entonces el tamaño de muestra n se puede calcular como

$$n = \frac{\frac{t^2 pq}{d^2}}{1 + \frac{1}{N} \left(\frac{t^2 pq}{d^2} - 1 \right)} \quad (3.20)$$

donde $q=1-p$ y t es la abscisa de la curva normal que corta un área de α en las colas de distribución.

Los tamaños de las muestras para cada estrato se distribuyeron proporcionalmente de acuerdo a la proporción que tiene cada estrato sobre la población total de 18 años en adelante; después de obtener el número de personas que se iba a entrevistar en cada estrato, se redistribuyó proporcionalmente sobre los municipios muestreados de cada estrato. En la tabla 1 se ilustra la distribución de las muestras y el tamaño de la muestra.

El operativo de campo consistió en reclutar a 85 encuestadores con grado de escolaridad pasante de licenciatura, los cuales fueron repartidos en 13 grupos; de éstos, siete se formaron con siete encuestadores y los otros seis con seis entrevistadores, se designaron tres grupos para cubrir León y diez para el resto de los estratos.

3.2 Resultados

3.2.1 Tamaño de Muestra

Sustituyendo en (3.20) los siguientes valores:

$$d=0.025 \quad \alpha=0.05 \quad t=1.96 \text{ (95\% de confianza)} \quad p=0.5 \quad N=1,944,161$$

obtenemos el tamaño de la muestra n para la encuesta preelectoral.

$$n=1,535.46 \cong 1,536$$

Se consideró un tamaño de muestra de 1536 respondientes más 68 adicionales, agregando en promedio un cuestionario por entrevistador, arrojando un total de 1604 encuestados.

GUANAJUATO							
DISTRIBUCIÓN MUESTRAL DE LA ENCUESTA PREELECTORAL							
Estrato	Rangos de los Tamaños de Municipios (Población 18 Años y Más)	Número de Municipios	Número de Municipios en Muestra	Población de 18 Años y Más	Población de 18 Años y Más Relativa %	Número de Encuestas	Proporción Relativa de Encuestas %
1	MENOS DE 10,000	12	6	72,054	3.61	66	4.11
2	(10,000 a 30,000)	16	12	333,489	16.72	269	16.77
3	(30,0001 a 70,000)	14	14	682,442	34.22	547	34.10
4	SALAMANCA	1	1	109,438	5.49	87	5.42
5	CÉLAYA	1	1	163,012	8.17	130	8.10
6	IRAPUATO	1	1	187,427	9.40	147	9.16
7	LEÓN	1	1	446,299	22.38	358	22.32
	TOTAL	46	36	1,994,161	100.00	1,604	100.00

Tabla 1. Distribución de la muestra por estratos encuesta preelectoral

3.2.2 Perfil Sociodemográfico de la Muestra

En esta sección presentaremos las diferentes variables sociales, demográficas y económicas que logró "fotografiar" la 2ª encuesta preelectoral levantada el 14 de mayo de 1995, por el Centro de Estudios de Opinión (CEO), de la Universidad de Guadalajara, a dos semanas de la

jornada electoral. Se hará una comparación (cuando sea posible) con el perfil sociodemográfico del INEGI basado en el XI Censo General de Población y Vivienda, 1990. Se debe recordar que la encuesta está considerando solamente a la población potencial con derecho a votar (18 años y más) tomada del código 90 del INEGI.

La primera variable es la edad, como es una variable cuota, se debe revisar que la partición de los grupos de edad designados por el CEO cumplan con la proporción correcta. Se verificó que de la población con derecho a voto, según código 90, el grupo de 18 a 24 años tiene una proporción de 27.3%; el de 25 a 34, un 26.3%; el de 35 a 44 años representa el 17.8% y el grupo de 45 y más con el 28.6%. Al comparar con la encuesta preelectoral, se aprecia que los primeros tres grupos sí cumplieron con sus cuotas correspondientes pero, el grupo de 45 y más resultó bajo, quizá esto se debió a que por el tipo de partición realizada, este grupo aglomeró un mayor número de edades y después no fue fácil cumplir con las cuotas que demandaba.

La otra variable cuota fue el sexo, la proporción de población masculina con derecho a voto según el INEGI es del 46.7%, sin embargo la proporción de hombres en la encuesta preelectoral es del 52.4%. Las figuras 1 y 2 presentan estas variables cuota respectivamente.

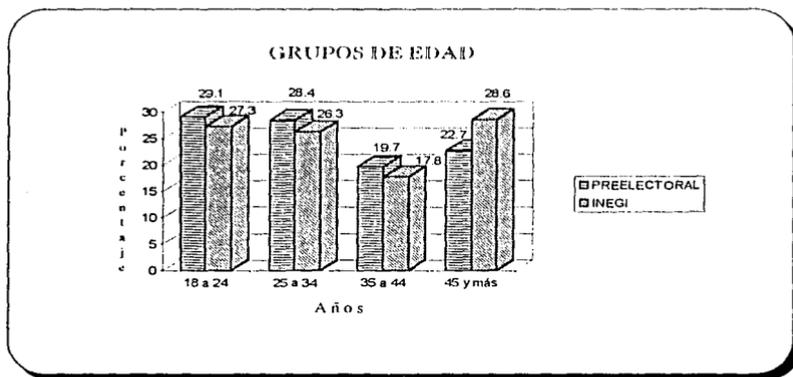


Figura 1. Grupos de edad encuesta preelectoral; 1575 respuestas válidas

La escolaridad en el estado, según la encuesta, refleja que entre primaria, secundaria y bachillerato esta repartida la mayor parte de la instrucción escolar de los encuestados. Se hará un comparativo gráfico aproximado, porque el INEGI maneja el nivel de instrucción solo en cuatro categorías que son: sin instrucción (19.1%); con primaria incompleta (26.2%); con primaria completa (21.0%) y con instrucción postprimaria (30.3%); aparte, la edad en la que desenvuelve su investigación es en la de 15 años y más. En la figura 3 se presentan las categorías de la encuesta preelectoral. En la figura 3A se presenta el comparativo con el perfil sociodemográfico

estatal del INEGI para el que se realizaron las siguientes reclasificaciones de categorías en ambas fuentes de información; en la encuesta preelectoral se fusionan las categorías de secundaria a posgrado bajo la categoría Postprimaria y en el INEGI se fusionan primaria incompleta y primaria completa bajo la categoría Primaria.

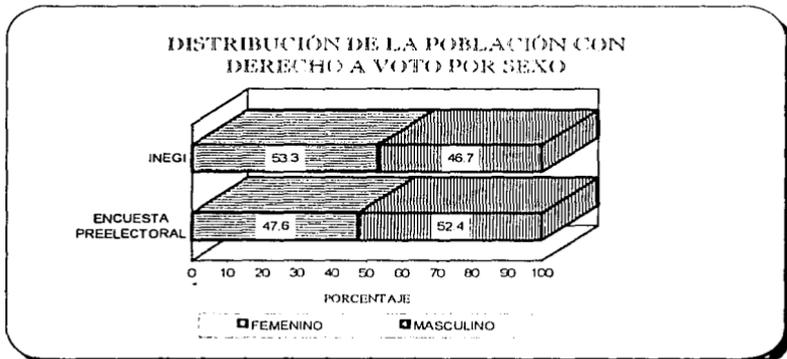


Figura 2. Distribución de la población con derecho a voto por sexo, 1582 respuestas válidas

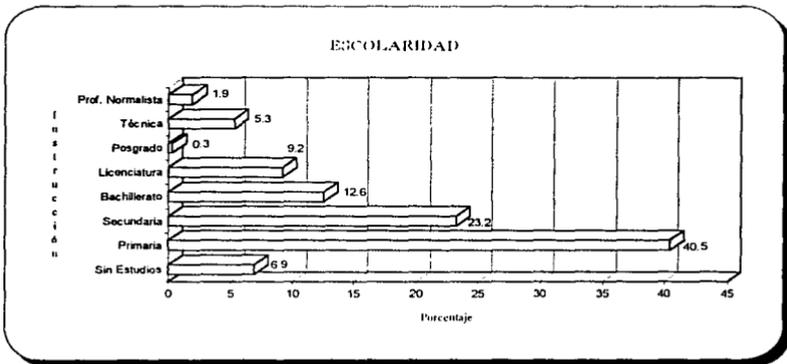


Figura 3. Nivel de instrucción encuesta preelectoral: 1507 respuestas válidas, 97 no especificadas. Total: 1604

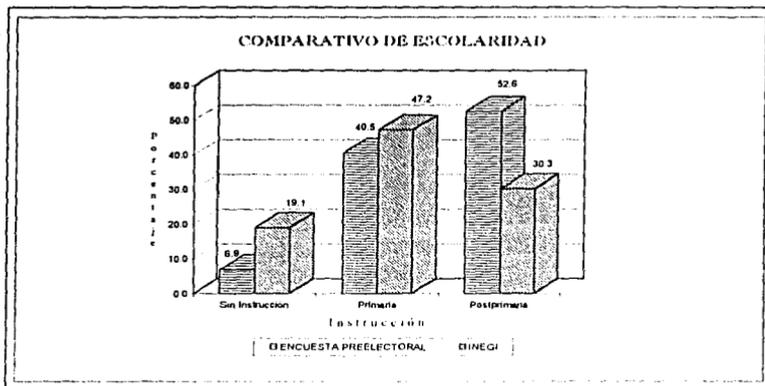


Figura 3A. Comparación de nivel de instrucción entre INEGI y Preelectoral; población utilizada por INEGI: 15 años y más

La variable ocupación, abarca una gran variedad de rubros; se observan serias diferencias con las que usa el perfil del INEGI, de entrada, las categorías que se podrían calificar de similares con las del INEGI están bajo otro concepto. Mientras que las categorías de la encuesta, como: ama de casa, estudiante, otro, etc., están como ocupación, el INEGI las coloca en *condición de actividad*, refiriéndose básicamente a la población económicamente activa e inactiva; otras categorías como: obrero, comerciante, etc., el INEGI las clasifica en *ocupación principal*, pero esta vez referenciada a que la ocupación está proporcionando un ingreso económico; por lo tanto, en la figura 4 sólo se muestran las clases que diseñó el CEO para la encuesta.

La última variable que se investigó en la encuesta fue el ingreso mensual. Aquí hubo mucha no respuesta, aunque quizá se justifica porque, obviamente, se tiene poca disposición para divulgar los ingresos personales y más si es en la vía pública. En la figura 5 se presentan las proporciones de cada rango diseñado por el CEO. El comparativo con el perfil sociodemográfico del INEGI se hará en la subsección 3.2.3A.

3.2.3 Asociación entre algunas Variables Manejadas en la Encuesta Preelectoral

En esta sección se realizarán algunos "cruces" de las diversas variables que posee la encuesta; hacer un "cruce" en el ámbito de la estadística quiere decir, que se observará cierta característica que asocia a dos categorías; por ejemplo, si se toma una de las categorías de la variable edad, digamos el grupo de 18 a 24 años y se desea saber qué proporción de este grupo votaría por un partido político en particular, se está haciendo un cruce, donde las categorías son el grupo de edad y el partido político, y la característica que las asocia es el voto.

Después de lo explicado en el párrafo anterior, se observará el comportamiento de las variables examinadas en la sección 3.2.2 cruzadas con otras variables llamadas electorales, que realmente son preguntas que no están relacionadas directamente con las preferencias políticas, sin embargo, fueron introducidas para medir y enfocar la cultura electoral que en ese momento se vivía en el estado. Por supuesto, también serán observados los cruces con las variables de preferencias, que, después de todo, son las que constituyen el objetivo fundamental de esta encuesta. Precisamente por lo anterior, esta sección se dividirá en dos subsecciones, la subsección 3.2.3A comprenderá los cruces de las variables electorales contra las variables sociodemográficas, y la subsección 3.2.3B contendrá los cruces de las variables de tendencias contra las sociodemográficas.

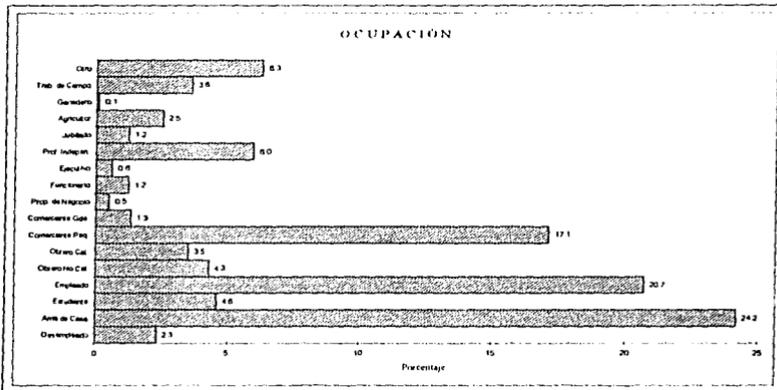


Figura 4. Ocupación: 1471 respuestas válidas, 133 no especificadas, Total 1604

3.2.3A Variables Electorales vs. Variables Sociales

Como se mencionó en la sección 3.2.3 la presente subsección está estructurada especialmente para analizar la relación entre el ambiente previo a los comicios y las distintas variables sociodemográficas extraídas por la misma encuesta; la estructura consistirá en: resaltar por escrito los valores máximos de cada cruce por categorías (que no se aprecian en las tablas), hacer algunos comentarios con respecto a éstos, ilustrar con gráficas y, proporcionar la información sobre el total a través de una tabla anexa a cada gráfica.

Antes de iniciar de lleno con las variables cuota se muestran algunos cruces que, además, rompen con lo establecido en este apartado, pero de alguna forma demuestran el grado de información electoral a dos semanas de la jornada. Y rompen con lo establecido porque estrictamente se trata de cruces de variables electorales.

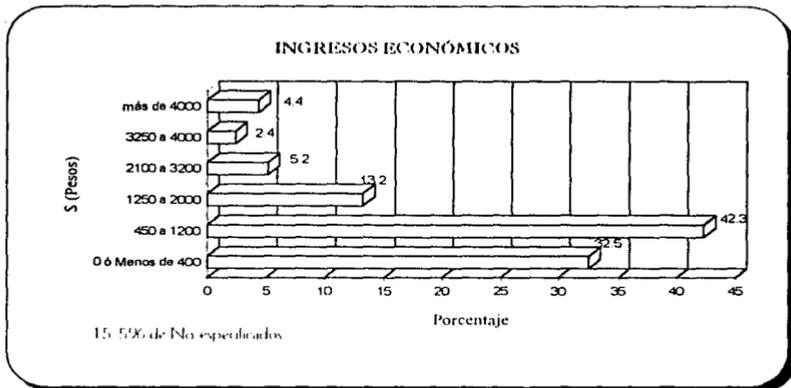


Figura 5. Ingresos registrados por encuesta preelectoral; 1356 respuestas válidas, 248 no especificados. Total 1604

Pues bien, el cruce es entre las preguntas: ¿Tiene credencial para votar? y ¿Piensa votar en próximas elecciones? Aquí se tienen dos aspectos por analizar, el primero es que había una clara intención de participar, porque de 1417 encuestados que sí tenían credencial de elector, 1317 contestaron que sí votarían, es decir, se trata del 92.9%, el segundo es que si a dos semanas de las elecciones una persona no posee credencial de elector, es obvio que no podrá votar aunque tenga la más firme intención de hacerlo, lo notable es que de los 130 encuestados que declararon no tener credencial, 43 contestaron que sí lo harían, 16 que aún no lo decidían y las restantes 71 la respuesta lógica puesto que no tienen credencial; entonces, algunas de las respuestas para explicar esto son: los encuestados no lo sabían; contestaron muy puntualmente (en ese preciso momento no tenían credencial); o, la más común en estos casos, no hubo suficiente información al respecto por parte de los organismos del gobierno responsables, sin embargo 59 personas (sumando 16 indecisas y 43 que sí lo harían) creían que podrían votar sin su credencial, sólo representan el 3.8%, esto podría ser un indicativo de que sí hubo una fuerte campaña informativa. La figura 6 muestra las respuestas válidas y la tabla 2 la información sobre el total.

Con el afán de continuar investigando sobre qué tan alto era el grado de cultura electoral, se realizó otro cruce cuyas preguntas fueron: ¿Piensa votar en próximas elecciones? y ¿Sabe la fecha de elecciones en Guanajuato? También se aprecia un alto grado de información, porque de 1359 respondientes que contestaron que sí pensaban votar en las elecciones, 1049 sabían exactamente la fecha en cuestión, lo cual representa el 77.2% de esta categoría, incluso de los 114 que contestaron negativamente el 60% sabían puntualmente el día. Se manejan otras respuestas cuyos resultados se presentan en la figura 7 y en la tabla 3 en el orden acostumbrado.

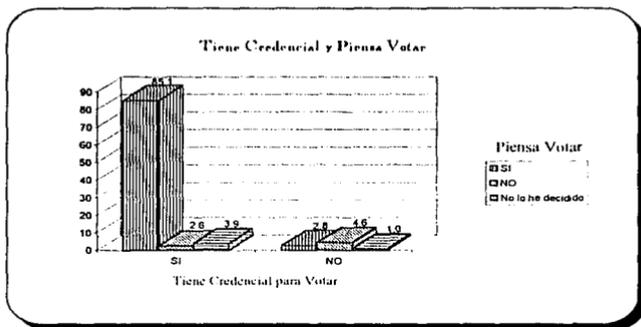


Figura 6. Cruce de credencial e intención de votar. 1547 repuestas válidas

¿Tiene credencial para votar?	¿Piensa votar en próximas elecciones?			Total
	SI	NO	No lo he decidido	
SI	85.1	2.6	3.9	91.6
NO	2.8	4.6	1.0	8.4
Total	87.9	7.2	4.9	100.0

Tabla 2. Cruce: credencial vs. intención de votar. 1547 respuestas

Para terminar este bloque de cruces de variables electorales se presenta el cruce consistente en la misma pregunta de ¿Piensa votar ...? y la pregunta ¿Sabe la ubicación de la casilla? Como en las anteriores observaciones, en ésta, gran parte de los respondientes que indicó que sí pensó votar, conocía la ubicación exacta de la casilla; quizá esto se debe a que ya se había distribuido el primero de los dos encartes⁶ que contenían la información sobre la ubicación de cada casilla en todo el estado, además de que la población estaba ansiosa por votar; el encarte advertía al elector que debía conservarlo porque sólo saldría ese día y el 27 de mayo de 1995 (un día antes de los comicios). Profundizando más sobre las categorías, se tiene que de 1349 personas que sí pensaban votar; 1085 sabían donde se ubicó la casilla que les correspondía, esto es el 80.4% que es alto para índices de participación ciudadana; por otra parte, de un total de 110 personas que no votarían, 47% sabían la ubicación y 52% no. Ver la figura 8. Sin embargo, la tabla 4 que

⁶ Encarte es un documento de aparición esporádica que se anexa a otro de circulación periódica, tal como un diario o una revista comúnmente de alto tiraje y amplio rango de distribución

contiene la información global, muestra un desconocimiento relativamente alto en cuanto a la ubicación de la casilla.

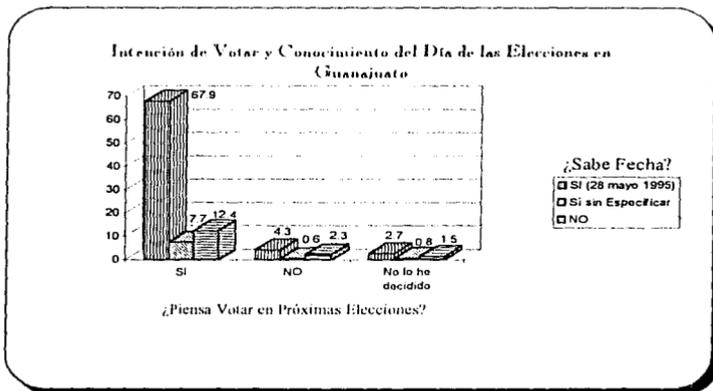


Figura 7. Cruce: ¿Piensa votar en próximas elecciones? y ¿Sabe fecha de elecciones en Guanajuato? 1545 respuestas válidas

¿Piensa votar en próximas elecciones?	¿Sabe fecha de elecciones en Guanajuato?			Total
	SI (28 mayo 1995)	SI sin Especificar	NO	
SI	67.9	7.7	12.4	88.0
NO	4.3	0.6	2.3	7.1
No lo he decidido	2.7	0.8	1.5	4.9
Total	74.8	9.1	16.1	100.0

Tabla 3. Cruce: intención de votar vs. conocimiento de la fecha de las elecciones. 1545 respuestas

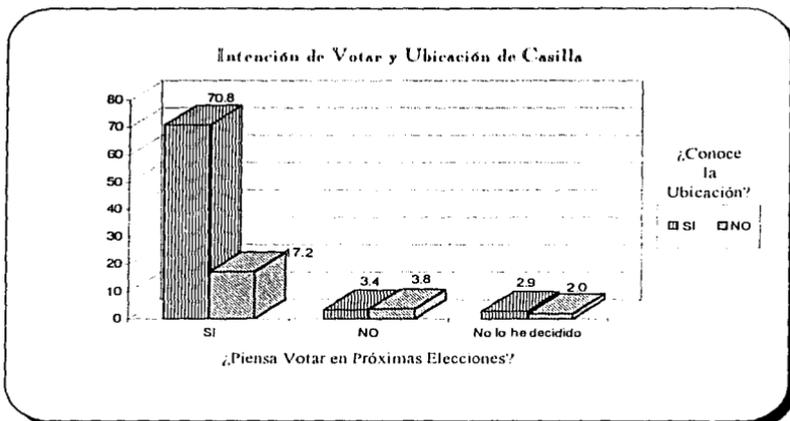


Figura 8. Cruce: ¿Piensa votar en próximas elecciones? y ¿Sabe la ubicación de la casilla?. 1533 respuestas validas

¿Piensa votar en próximas elecciones?	¿Sabe la ubicación de la casilla?		Total
	SI	NO	
SI	70.8	17.2	88.0
NO	3.4	3.8	7.2
No lo he decidido	2.9	2.0	4.8
Total	77.0	23.0	100.0

Tabla 4. Cruce: intención de votar vs. ubicación de la casilla correspondiente

A continuación se trabajará con las variables cuota, la primera de éstas es la edad cuyos cruces son contra: la credencial, la intención de votar, la fecha de las elecciones y la ubicación de la casilla. Cabe aclarar que, por ser variable cuota dependiente de las proporciones de la distribución poblacional, el grupo de edad dominante, como se vio en la sección 3.2.2, es el de 18 a 24 años, seguido muy de cerca por el de 25 a 34; por lo tanto, en los cruces serán los que resalten más y por tal motivo no se puede decir que fueron los más participativos, por esto mismo, se tratan a los grupos en forma independiente, para ver dentro de cada uno el grado de

información y de participación en las votaciones, aunque como los dos grupos anteriormente mencionados varían por 11 personas en cuanto al tamaño, tal vez esto permitirá hacer algunas acotaciones relacionando a ambos.

El primer cruce es el de la edad y la posesión o carencia de la credencial de elector; de las 459 personas que tienen entre 18 y 24 años; 67 no poseen credencial, representan el 14.6% pero, aplicando esta tasa al grupo de edad 25 a 34 que consta de 445 personas, da un resultado aproximado de 65 personas que no tienen credencial, sin embargo, solamente 29 entrevistados pertenecientes a este grupo dijeron no tenerla, por lo tanto se puede decir que el grupo de 25-34 está más conscientizado sobre el deber de votar, con el simple hecho de tener su credencial de elector. En los otros dos grupos se notan tasas relativamente bajas sobre el mismo análisis; la gráfica correspondiente a esta observación se encuentra en la figura 9 y, las proporciones sobre el total de las observaciones, en la tabla 5.

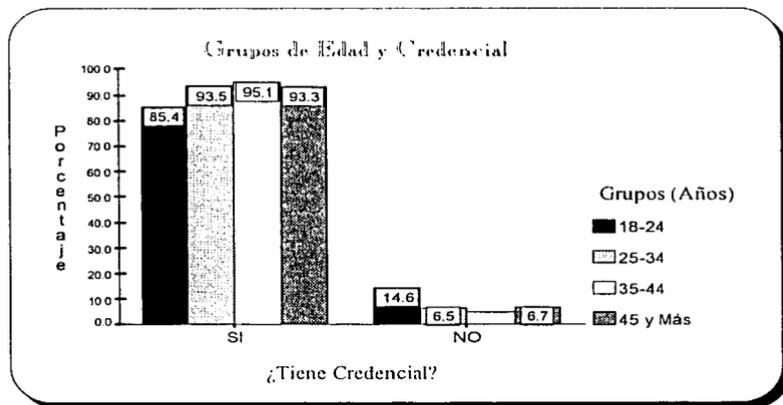


Figura 9. Cruce: Edad y Tenencia de credencial. 1570 respuestas válidas

Para el cruce de la edad y la intención de votar, se aprecia que prevalece el mismo comportamiento para los grupos 18-24 y 25-34, a simple vista, existe más vocación ciudadana en el grupo de edad 25-34; 391 de los respondientes indicaron que sí votarían y, en cambio, en el grupo precedente contestaron en forma afirmativa solamente 379 encuestados. En los otros dos grupos, también vemos una alta intención de votar. Para cerrar con este cruce, no hay que perder de vista que en total 108 respondientes manifestaron que no votarán, esto es aproximadamente el 7% de las 1526 respuestas válidas de este cruce. La figura 10 demuestra que el grupo 25-34 es el mayor en cuanto a participación y la tabla 6 hace lo propio con la información global.

GRUPOS DE EDAD (Años)	¿Tiene credencial de elector?		Total Renglón
	SI	NO	
18 - 24	25.0	4.3	29.3
25 - 34	26.5	1.8	28.3
35 - 44	18.7	1.0	19.7
45 y más	21.2	1.5	22.7
Total	91.4	8.6	100.0

Tabla 5. Cruce: Edad vs. credencial. Preelectoral 14 mayo 1995. 1570 respuestas

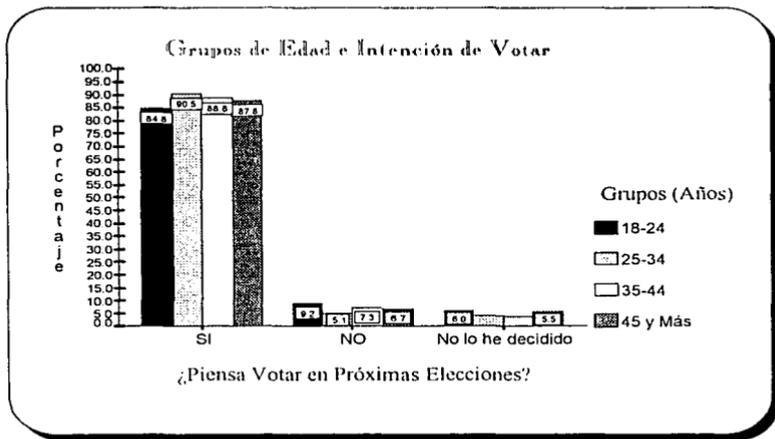


Figura 10. Cruce: Edad e intención de votar. 1526 respuestas válidas

GRUPOS DE EDAD (Años)	¿Piensa votar en próximas elecciones?			Total Renglón
	SI	NO	No lo he decidido	
18 - 24	24.8	2.7	1.8	29.3
25 - 34	25.6	1.4	1.2	28.3
35 - 44	17.6	1.4	0.8	19.9
45 y más	19.8	1.5	1.2	22.5
Total	87.9	7.1	5.0	100.0

Tabla 6. Cruce: Edad vs. intención de votar. 1526 respuestas

Continuando con la edad, en seguida se analizará, el cruce con el conocimiento de la fecha. Una vez más existe la misma tendencia; hasta este momento se han hecho tres cruces y poco a poco empieza a notarse que de los dos grupos que podrían considerarse jóvenes, el grupo 25 a 34 años sobresale; por lo que se harán algunos comentarios más adelante. Por lo pronto, la tabla 7 presenta una proporción de ignorancia de la fecha del 16% global, y en la figura 11 la categoría 18-24 es la que tiene la tasa más alta de ignorancia.

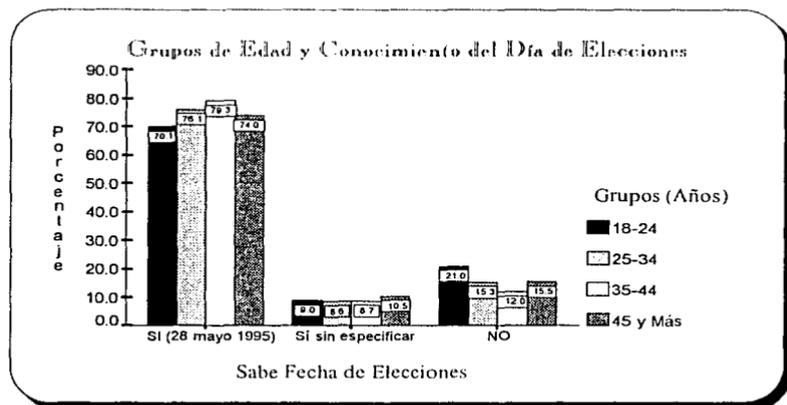


Figura 11. Cruce: Edad y fecha de elecciones. 1565 respuestas válidas

GRUPOS DE EDAD (Años)	¿Sabe la fecha de elecciones en Guanajuato?			Total Rengión
	SI (28 mayo 1995)	Si sin Especificar	NO	
18 - 24	20.5	2.6	6.1	29.3
25 - 34	21.6	2.4	4.3	28.4
35 - 44	15.7	1.7	2.4	19.7
45 y más	16.7	2.4	3.5	22.6
Total	74.5	9.1	16.4	100.0

Tabla 7. Cruce: Edad vs. conocimiento de la fecha de las elecciones. 1565

El último cruce de la variable edad se hizo contra la ubicación de la casilla (figura 12 y tabla 8), vuelven a notarse algunas similitudes con referencia a los cruces anteriores pero, aquí se refleja en qué edades hubo poca información al respecto porque, el 23% del total global ignora la ubicación de la casilla; por otro lado este es el único cruce donde no sobresalió el grupo 25-34 años.

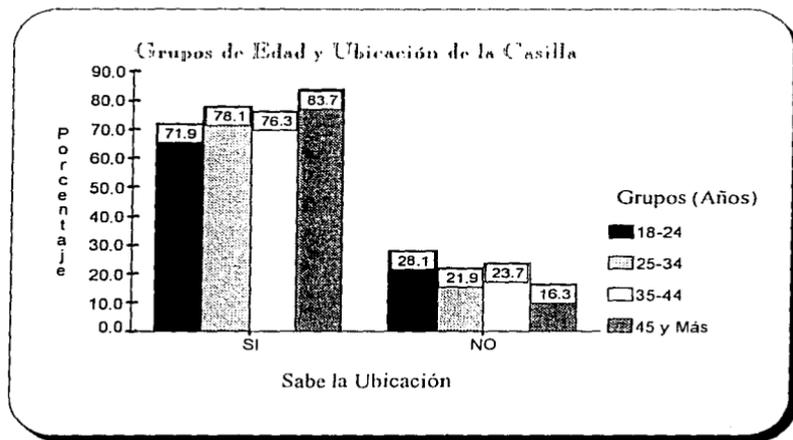


Figura 12. Cruce: Edad y ubicación de la casilla correspondiente. 1549 respuestas válidas

GRUPOS DE EDAD (Años)	¿Sabe la ubicación de la casilla?		Total Renglón
	SI	NO	
18 - 24	20.9	8.1	29.0
25 - 34	22.3	6.3	28.5
35 - 44	15.2	4.7	19.9
45 y más	18.9	3.7	22.6
Total	77.2	22.8	100.0

Tabla 8. Cruce: Edad vs. ubicación de casilla. 1549 respuestas

Después, de analizar los cruces de la primera variable cuota. Es muy claro que el grupo que más se notó fue el de 25-34 años, incluso siendo menor en 11 personas al de 18-24 años. Es muy difícil saber porqué este grupo sobresalió pero, es muy común que el individuo en esa etapa de la vida esta en un nivel muy alto de responsabilidad en todos sus aspectos; por ejemplo, tiene hijos y piensa en que les puede dejar, en consecuencia, siente un gran impulso por contribuir en la comunidad, y si cree que participando en ésta puede lograr un cambio, pues no va a dudar en formar parte del cambio, por lo tanto, se motiva y adquiere un potencial muy alto para llevar a cabo sus ideales. Por otro lado hay que considerar que se trataba de elecciones extraordinarias, es decir, se entraba a un proceso de volver a elegir con más cuidado y con la posibilidad de lograr un cambio histórico.

Para la segunda variable cuota se realizaron los mismos cruces; así que primero se presentan las gráficas y las tablas respectivas; los comentarios se harán al final.

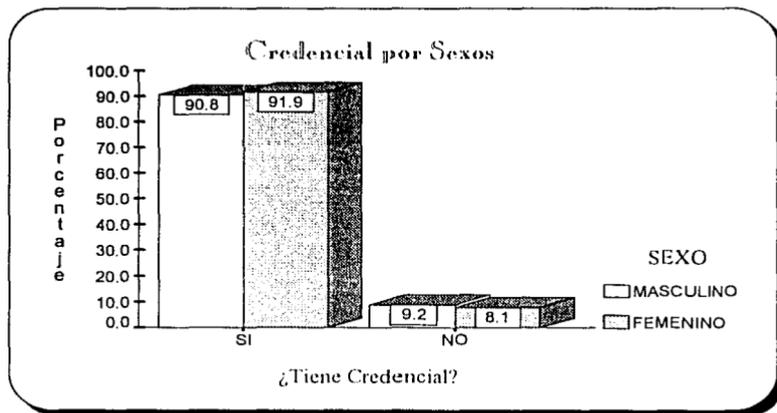


Figura 13. Cruce: Sexo y credencial. 1577 respuestas válidas

SEXO	¿Tiene credencial de elector?		
	SI	NO	Suma Renglón
MASCULINO	47.5	4.8	52.3
FEMENINO	43.8	3.9	47.7
Suma Columna	91.3	8.7	100.0

Tabla 9. Cruce: Sexo vs. Credencial. 1577 respuestas

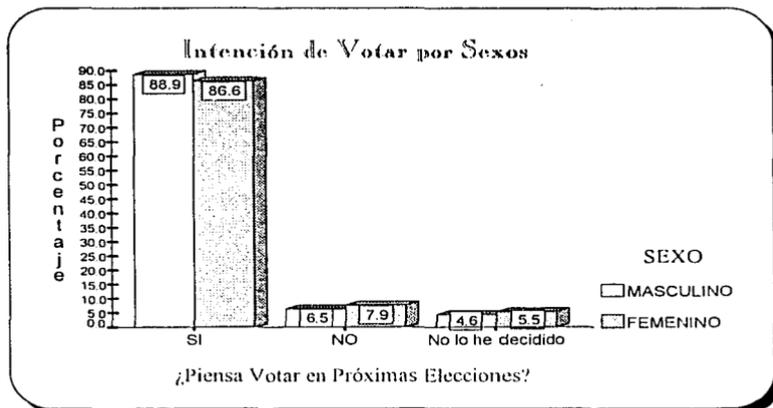


Figura 14. Cruce: Sexo y Piensa votar. 1529 respuestas válidas

SEXO	¿Piensa votar en próximas elecciones?			Suma Renglón
	SI	NO	No lo he decidido	
MASCULINO	46.8	3.4	2.4	52.6
FEMENINO	41.1	3.7	2.6	47.4
Suma Columna	87.9	7.1	5.0	100.0

Tabla 10. Cruce: Sexo vs. Intención de votar. 1529 respuestas

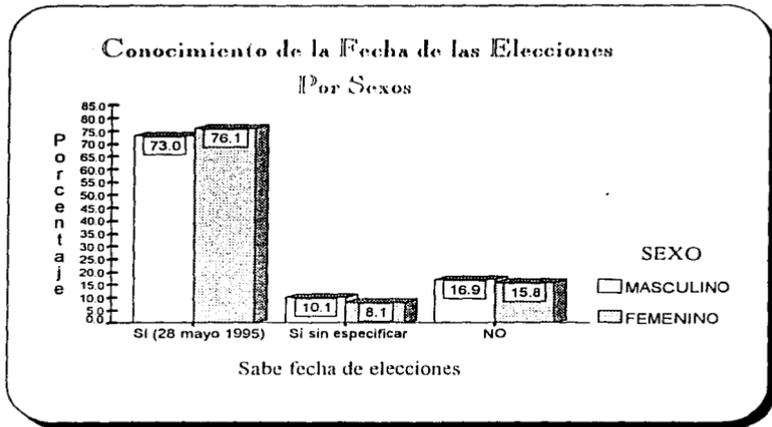


Figura 15. Cruce: Sexo y Pienso votar. 1572 respuestas válidas

		¿Sabe fecha de elecciones en Guanajuato?			
S E X O	SI (28 mayo 1995)	Si sin Especificar	NO	Suma Renglón	
MASCULINO	38.2	5.3	8.8	52.3	
FEMENINO	36.3	3.9	7.5	47.7	
Suma Columna	74.5	9.2	16.3	100.0	

Tabla 11. Cruce: Sexo vs. Conocimiento de la fecha de las elecciones. 1572 respuestas

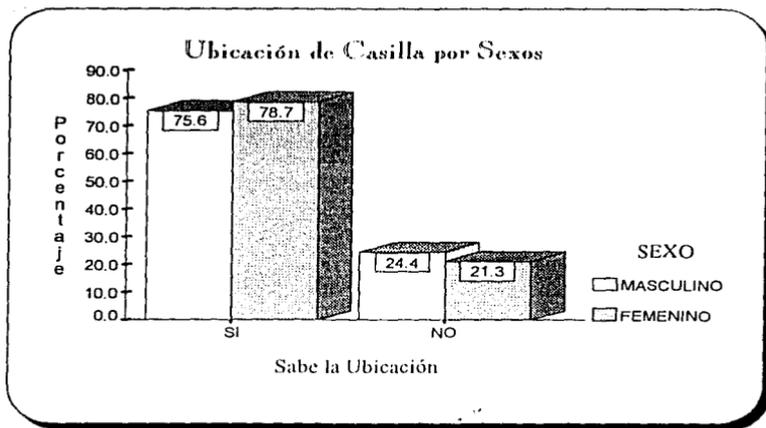


Figura 16. Cruce: Sexo y Ubicación de la casilla correspondiente. 1554 respuestas

SEXO	¿Sabe la ubicación de la casilla?		
	SI	NO	Suma Renglón
MASCULINO	39.7	12.8	52.5
FEMENINO	37.4	10.1	47.5
Suma Columna	77.1	22.9	100.0

Tabla 12. Cruce: Sexo vs. Ubicación de casilla. 1554 respuestas

El primer análisis es ver qué proporción de mujeres y hombres hay sobre las respuestas válidas a la pregunta con relación a la credencial de elector; hubo en total 1577 personas, de las cuales 825 son hombres y 752 mujeres, en proporciones de 52.3% y 47.6% respectivamente. Ahora se verá por grupos cuántos individuos tienen credencial: 691 mujeres tienen credencial, es decir, casi el 92%; los hombres con credencial son 749, esto da casi el 91%. La diferencia es

pequeña, lo cual indica que ambos sexos tenían interés por participar al contar con su credencial. Los resultados se muestran gráficamente en la figura 13 y los totales en la tabla 9.

En relación a si piensan votar, tenemos un total de 1529 respuestas válidas el 52.6% (804) para hombres y el 47.4% (725) para mujeres. Dentro del grupo femenino, 628 si van a votar (86.6%), 40 están indecisas (5.5%) y 57 no lo harán (7.9%); para el masculino en el mismo orden, 715 (88.9%), 52 (6.5%) y 37 (4.6%). En este cruce, también se ven cifras parecidas entre sexos. Información en figura 14 y proporciones sobre el total en la tabla 10.

Para la fecha de las elecciones, existen 1572 respuestas válidas; 749 mujeres representan el 47.6% y 823 hombres el 52.4%. En el grupo masculino 601 contestaron exactamente (73.0%), 83 no supieron la fecha exacta (10.1%) y 139 la ignoraban (16.9%); del grupo femenino, sabían la fecha exacta 570 (76.1%), no la mencionaron 61 (8.1%) y la ignoraban 118 (15.8%). Aquí se nota un 3% de diferencia entre los grupos en cuanto al conocimiento de la fecha exacta. Gráfica en la figura 15 y proporciones globales en la tabla 11.

Por último, para la ubicación de la casilla, se obtuvieron 1554 respuestas válidas; de las cuales 816 son hombres y 738 mujeres, cada grupo representa el 52.5% y el 47.5% respectivamente. Las mujeres que sabían la ubicación de la casilla eran 581 esto es el 78.7% y las que no lo sabían fueron 157 el 21.3%. Los hombres que ubicaban la casilla fueron 617 (75.6%) y el complemento, 199 (24.4%). Otra diferencia del 3% a favor del sexo femenino.

Haciendo un comentario global sobre esta segunda variable cuota, se notó tal semejanza en las cifras que sería muy difícil hacer cualquier comentario referido a distinción entre sexos.

Después de las variables cuota, se cruzarán la escolaridad y dos preguntas de la información electoral, éstas son la de la credencial y la de intención de votar. Para efecto de estos cruces, las categorías originales de la encuestase reducirán (figura 3 de este capítulo) de ocho a cinco. Estas modificaciones son: Sin Estudios, Primaria y Secundaria quedaran intactas; Bachillerato y Técnica quedan bajo el nombre de Bachillerato, y, Licenciatura, Profesor Normalista y Posgrado se unifican para crear la categoría Postbachillerato.

En la figura 17 y en la tabla 13 se observa que los grupos de más peso (Primaria y Secundaria), tienen una proporción muy alta en relación a la posesión de la credencial de elector, haciendo una observación más profunda se nota que los grupos restantes sobrepasan el 90%, como consecuencia tenemos una proporción global del 91.3% de personas que sí cuentan con la credencial.

Son tan similares las proporciones de los grupos de escolaridad de la figura 18, que no tendría ningún sentido realizar algún comentario comparativo.

La tabla 14 confirma lo anterior porque del total de las respuestas, el 88% si piensa votar. Para dar una idea del alto grado de la intención de participar, supongamos que se elimina el grupo escolar más grande que es Primaria de la columna del SI y que se distribuye en las otras dos respuestas, aún así se tiene un índice de participación aceptable.

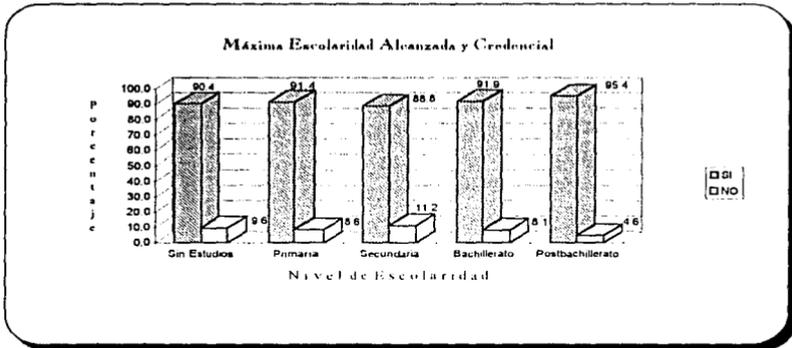


Figura 17. Cruce: Escolaridad y Credencial. 1504 respuestas válidas

ESCOLARIDAD	¿Tiene credencial de elector?		Suma Renglón
	SI	NO	
SIN ESTUDIOS	6.3	0.7	6.9
PRIMARIA	37.0	3.5	40.4
SECUNDARIA	20.6	2.6	23.2
BACHILLERATO	16.5	1.5	18.0
POSTBACHILLERATO	11.0	0.5	11.5
Suma Columna	91.3	8.7	100.0

Tabla 13. Cruce: Escolaridad vs. Credencial. 1504 respuestas

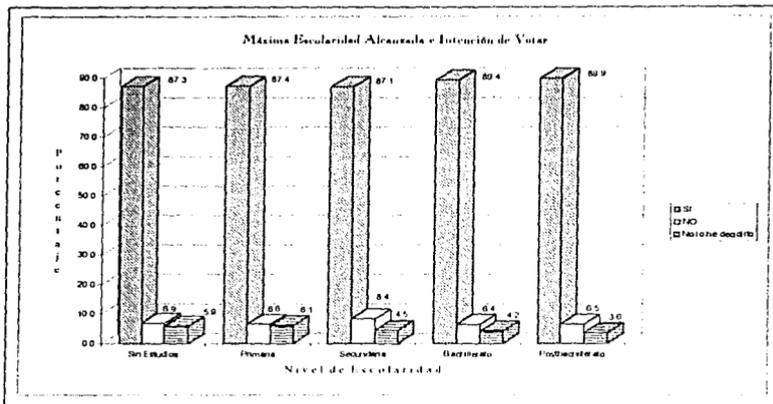


Figura 18. Cruce: Escolaridad e intención de votar. 1461 respuestas válidas

ESCOLARIDAD	¿Piensa votar en próximas elecciones?			Suma Renglón
	SI	NO	No lo he decidido	
SIN ESTUDIOS	6.1	0.5	0.4	7.0
PRIMARIA	35.5	2.7	2.5	40.6
SECUNDARIA	19.9	1.9	1.0	22.9
BACHILLERATO	16.2	1.2	0.8	18.1
POSTBACHILLERATO	10.3	0.8	0.4	11.5
Suma Columna	88.0	7.0	5.1	100.0

Tabla 14. Cruce: Escolaridad vs. intención de votar. Categorías modificadas. 1461 respuestas

Continuando con la ocupación, esta variable posee 17 categorías de las cuales algunas son prácticamente insignificantes, como se puede apreciar en la figura 4 de este capítulo, por este motivo también se realizarán algunas modificaciones a las categorías. Cabe recordar que tomando la variable individualmente, se tiene una proporción del 8,3% de no especificados, y que al cruzarla, posiblemente aumentará. Se practicarán los mismos cruces que se hicieron con la escolaridad.

Los cambios son los siguientes: desempleado y jubilado se convierten en la categoría Desempleado o Jubilado; obrero calificado y no calificado quedan como Obrero; las dos categorías de comerciante y propietario de negocio se fusionan para crear Comerciante; funcionario y profesionista independiente se unen a Ejecutivo y quedan con el nombre de este último; finalmente, agricultor, ganadero y trabajador de campo forman la clasificación Agropecuario. El resto de los grupos permanecen igual; así entonces las clasificaciones de la variable ocupación se reducen de 17 a 9 y quedan identificadas en la figura 19.

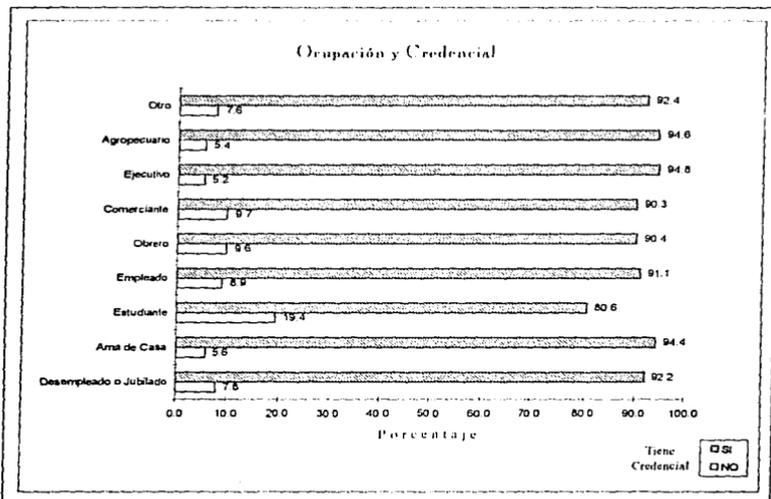


Figura 19. Cruce: Ocupación y Credencial. 1468 respuestas válidas.

OCUPACIÓN	¿Tiene credencial de elector?		
	SI	NO	Suma Renglón
Desempleado o Jubilado	3.2	0.3	3.5
Ama de Casa	22.9	1.4	24.3
Estudiante	3.7	0.9	4.6
Empleado	18.9	1.8	20.7
Obrero	7.0	0.7	7.8
Comerciante	17.0	1.8	18.9
Ejecutivo	7.4	0.4	7.8
Agropecuario	5.9	0.3	6.3
Otro	5.8	0.5	6.3
Suma Columna	91.8	8.2	100.0

Tabla 15. Cruce: Ocupación vs. Credencial. 1468

Bajo el análisis de la variable ocupación, en la figura 19 se observa que la categoría estudiante tiene la tasa más baja de tenencia de credencial, esto puede deberse a diversos factores: uno es que los estudiantes apenas están entrando a la mayoría de edad y aún no tienen la conciencia cívica de ejercer el sufragio, por lo tanto, no consideran importante poseer la credencial; otro es que si por algún motivo repitieron el trámite para hacerse de la credencial, tal vez no les fue posible obtenerla antes de las elecciones. Ahora bien, a nivel individual el 80% no es una tasa baja, pero, en comparación con los demás grupos cuyas tasas se encuentran entre el 90 y 95%, se hace necesario buscar una explicación al porqué de la diferencia con los demás.

La tabla 15 muestra las proporciones de cada clase en general; las proporciones de la respuesta NO son muy pequeñas, esto es porque el 8.2% del total de la encuesta no posee credencial. Los grupos de más peso son: Ama de Casa, Empleado y Comerciante, estos grupos en conjunto representan el 64% del total de la encuesta.

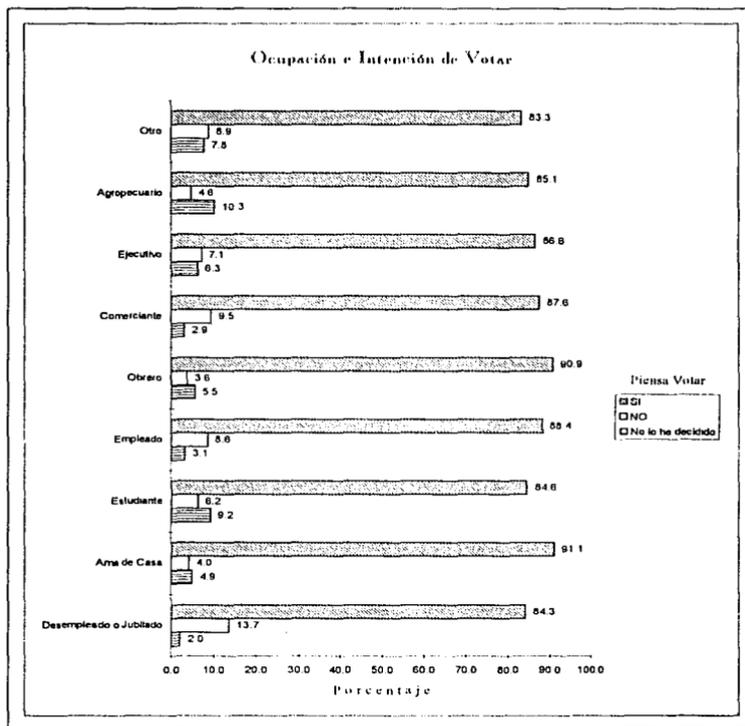


Figura 20. Cruce: Ocupación e Intención de Votar. 1.431 respuestas válidas

OCUPACIÓN	¿Piensa votar en próximas elecciones?			
	SI	NO	No lo he decidido	Suma Renglón
Desempleado o Jubilado	3.0	0.5	0.1	3.6
Ama de Casa	22.3	1.0	1.2	24.5
Estudiante	3.8	0.3	0.4	4.5
Empleado	18.0	1.7	0.6	20.4
Obrero	7.0	0.3	0.4	7.7
Comerciante	16.8	1.8	0.6	19.1
Ejecutivo	6.8	0.6	0.5	7.9
Agropecuario	5.2	0.3	0.6	6.1
Otro	5.2	0.6	0.5	6.3
Suma Columna	88.1	7.0	4.9	100.0

Tabla 16. Cruce: Ocupación vs. Intención de votar. 1431 respuestas

En las otras variables que se cruzaron con la intención de votar, se observaron valores aproximados entre el 85 y 90 por ciento para la respuesta afirmativa; el cruce que muestra la figura 20 (ocupación y votar) no fue la excepción, salvo la clasificación Otro, que se salió ligeramente del rango citado. La tabla 16 muestra que el 88% de 1431 respondientes si piensan votar y el resto, no lo harán o aún no lo deciden. La proporción de no especificados en este cruce fue aproximadamente del 10.8%.

Para esta variable es difícil hacer comentarios, porque como se había mencionado en la sección 3.2.2, involucra problemas de tipo conceptual al tratar de clasificar las ocupaciones de la población, ya que tiene como ocupaciones categorías que realmente no proporcionan un ingreso, y ocupación siempre se relaciona de inmediato con trabajo remunerado. Sin embargo, muestra un comportamiento semejante a los otros cruces y una clara intención de participar en los comicios.

Para finalizar esta subsección se presentan los cruces de la variable ingresos contra la credencial, (figura 21 y tabla 17) y la intención de votar (figura 22 y tabla 18).

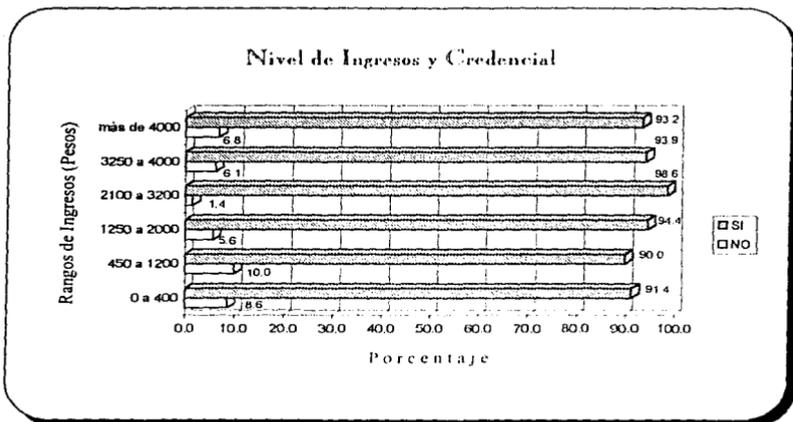


Figura 21 Cruce: Ingresos y Credencial. 1353 respuestas válidas

INGRESOS \$ Pesos	¿Tiene credencial de elector?		
	SI	NO	Suma Renglón
0 a 400	29.7	2.8	32.5
450 a 1200	38.1	4.2	42.3
1250 a 2000	12.5	0.7	13.2
2100 a 3200	5.1	0.1	5.2
3250 a 4000	2.3	0.1	2.4
más de 4000	4.1	0.3	4.4
Suma Columna	91.7	8.3	100.0

Tabla 17. Cruce: Ingreso vs. Credencial. 1353 respuestas

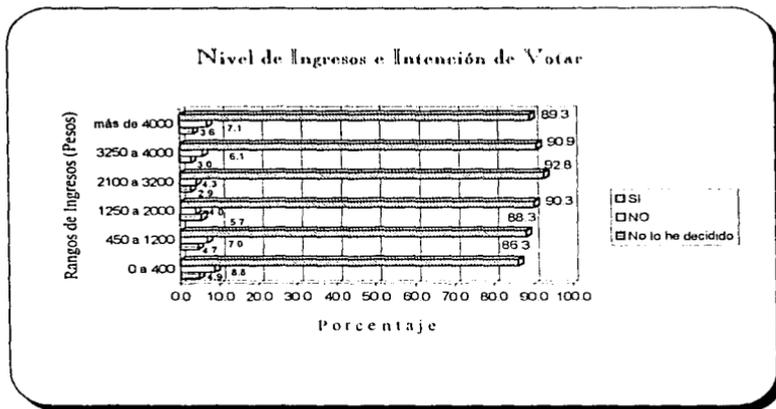


Figura 22 Cruce: ingresos e intención de Votar. 1319 respuestas válidas

INGRESOS \$ Pesos	¿Piensa votar en próximas elecciones?			Suma Renglón
	SI	NO	No lo he decidido	
0 a 400	28.1	2.9	1.6	32.6
450 a 1200	37.2	3.0	2.0	42.2
1250 a 2000	12.0	0.5	0.8	13.3
2100 a 3200	4.9	0.2	0.2	5.2
3250 a 4000	2.3	0.2	0.1	2.5
más de 4000	3.8	0.3	0.2	4.2
Suma Columna	88.2	7.1	4.7	100.0

Tabla 18. Cruce: Ingreso vs. Intención de votar. 1319 respuestas

Después de lo exhaustivo que ha sido esta subsección no hay mucho que comentar sobre este último cruce, porque presenta patrones de proporcionalidad similares a las variables anteriores; sin embargo, refleja de alguna forma la desconfianza que tienen las personas cuando se les cuestiona acerca de sus ingresos monetarios. Fue la variable con mayor índice de no especificación en cada uno de sus cruces, el primer cruce tuvo el 15.6% y el segundo 17.7%. Haciendo una comparación aproximada con la información del perfil sociodemográfico del estado publicado por el INEGI, los tres primeros grupos tienen un comportamiento parecido. Esta comparación resulta burda, porque el INEGI maneja las categorías por S.M. (salario mínimo⁷), las cuales, con sus respectivas proporciones son: *No recibe Ingresos* 8.0%; *Menos de un S.M.* 17.6%; *De 1 a 2 S.M.* 35.8%; *Más de 2 y menos de 3 S.M.* 15.0%; *De 3 a 5 S.M.* 10.5%; *Más de 5 S.M.* 7.5% y *No especificado* 5.6%. Estas proporciones están calculadas sobre 1,030,160 guanajuatenses.

En términos de rangos de pesos con la reforma monetaria tal como lo maneja la encuesta serían respectivamente: \$0.0; menos de \$252.15; de \$252.15 a \$504.30; más de \$504.30 y menos de \$756.45; de \$756.45 a \$1260.75 y más de \$1260.75. Es evidente que los rangos en pesos actuales diseñados por el INEGI se traslapan con los de la encuesta, pero se pueden fusionar algunas categorías del INEGI para hacerlas equivalentes a las del CEO y así lograr una comparación refinada.

En la encuesta preelectoral los tres primeros rangos reúnen el 88% de total de la muestra mientras que en el INEGI las cuatro primeras clasificaciones suman el 76.4% del total.

3.2.3B Preferencias Partidarias vs. Variables Sociodemográficas

Esta subsección trata los resultados de lo que es el alma de esta encuesta, en la tabla 19, que contiene las respuestas a la pregunta: Si hoy fueran las elecciones para decidir al próximo gobernador del estado de Guanajuato, ¿Por cuál partido votaría?. La encuesta preelectoral consideró como principales fuerzas políticas a cinco partidos que son: Partido Acción Nacional (PAN), Partido Revolucionario Institucional (PRI), Partido de la Revolución Democrática (PRD), Partido del Trabajo (PT) y Partido Frente Cardenista de Reconstrucción Nacional (PFCRN); a los restantes, los agrupó en la categoría etiquetada con OTRO, éstos son: Partido Auténtico de la Revolución Mexicana (PARM), Partido Demócrata Mexicano (PDM), Partido Popular Socialista (PPS) y Partido Verde Ecologista de México (PVEM). La tabla presenta valores que no se obtuvieron con el diseño muestral y los resultados oficiales de la contienda electoral.

La tabla 19 proporciona un comparativo rápido de tendencias entre la encuesta y los resultados oficiales. La primera observación es que coinciden en cuanto a los lugares que ocupó cada partido en la votación pero, es necesario aclarar algunos aspectos: el primero es que los cálculos de la preelectoral están hechos sobre 1548 respuestas válidas, es decir, no se consideran no especificados; segundo, existen dos respuestas que provocan desviación a la preferencia por un partido, éstas son la de la indecisión, que tiene un porcentaje significativo y la de

⁷ Hasta el 12 de marzo de 1990, el salario mínimo vigente en Guanajuato fue de \$252,150 pesos viejos mensuales, con la reforma monetaria actual (1996) son \$252.15 pesos.

simplemente no votar, cuyo porcentaje no es importante, los resultados oficiales del IEE Guanajuato no incluían a los otros partidos políticos.

PARTIDO	Encuesta Preelectoral	Resultados Oficiales ¹
	14 / mayo / 95	28 / mayo / 95
PAN	50.9%	58.1%
PRI	25.0%	32.9%
PRD	6.5%	7%
PT	1.1%	1.2%
PFCRN	0.3%	0.8%
OTRO	0.5%
NINGUNO	3.2%
NO HE DECIDIDO	12.5%
TOTAL	100.0	100.0
¹ Fuente: IEE de Guanajuato, proporciones sobre 1,944,955 votos válidos		

Tabla 19. Proporciones de las preferencias partidarias sin diseño de muestreo. 1548 respuestas

A continuación se describe la forma en que se analizará en esta subsección. El objetivo de estos análisis, es el de mostrar cómo un cruce por columnas no es igual a uno por renglones convirtiéndose en otra manera de interpretar cada cruce.

Este análisis se presentará con las siguientes versiones:

Análisis 1.

Gráficas basadas exclusivamente en las frecuencias acumuladas por los tres principales partidos políticos; asimismo, mostrarán las proporciones alcanzadas por cada uno dentro de las categorías de la variable sociodemográfica en turno. Al pie de cada figura se dan las respuestas y la proporción que representan del total. Cada figura está asociada a una tabla que contiene las respuestas totales por columna (categoría).

Análisis 2.

Se despliegan tablas equivalentes a las del análisis 1, donde la equivalencia consiste en que éstas contendrán las proporciones de las respuestas totales por renglón (partido); es decir, se presentan las proporciones alcanzadas por cada categoría de la variable sociodemográfica en cuestión dentro de cada partido.

Para lograr una clara relación entre las tablas de ambos procesos analíticos, las del análisis 2 se identifican con el mismo número que su similar en el análisis 1 más la letra "A".

ANÁLISIS I

La figura 23 muestra que el PAN domina en la mayoría de los grupos a un razón de dos a uno con respecto a el PRI, con excepción del grupo 45 y más que no cumple con esta razón, el PRD tiene valores poco significativos, sin embargo los votos que recibió se repartieron casi en forma equitativa para cada grupo de edad.

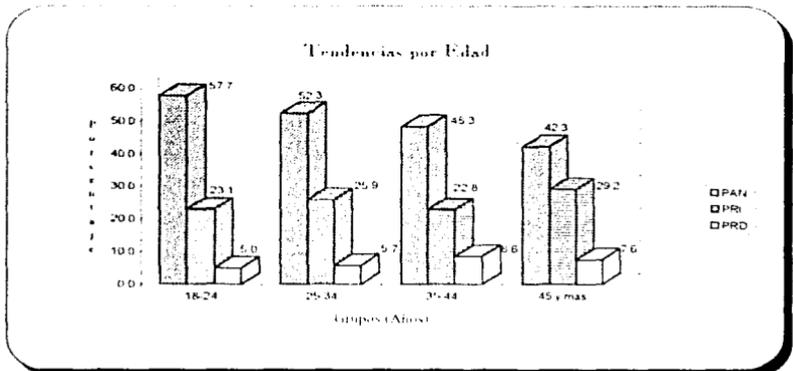


Figura 23 Cruce: Preferencias vs. Edad. 1257 respuestas acumuladas. 82.5% de 1523

PARTIDO	GRUPOS DE EDAD (AÑOS)			
	18 - 24	25 - 34	35 - 44	45 y más
PAN	57.7	52.3	48.3	42.3
PRI	23.1	25.9	22.8	29.2
PRD	5.0	5.7	8.6	7.6
PT	0.7	1.8	0.7	1.2
PFCRN	0.5	0.2	0.0	0.6
OTRO	0.2	0.9	0.3	0.6
Ninguno	2.7	2.8	3.6	4.1
No he Decidido	10.2	10.3	15.6	14.6
Suma	100.0	100.0	100.0	100.0
Frecuencias (n)	442.0	436.0	302.0	343.0

Tabla 20. Cruce: Preferencias partidarias vs. Grupos de edad. 1523 respuestas

La tabla 20 contiene las proporciones de las tendencias partidarias por grupos de edad. En esta tabla se despliegan todas las categorías con el fin de tener una dimensión de las proporciones de cada opción partidista; en las subsecuentes aparecerán los porcentajes del PAN, PRI, PRD, y se acumulará en Otro las proporciones del PT, PFCRN y OTRO; las opciones restantes no se alteran.

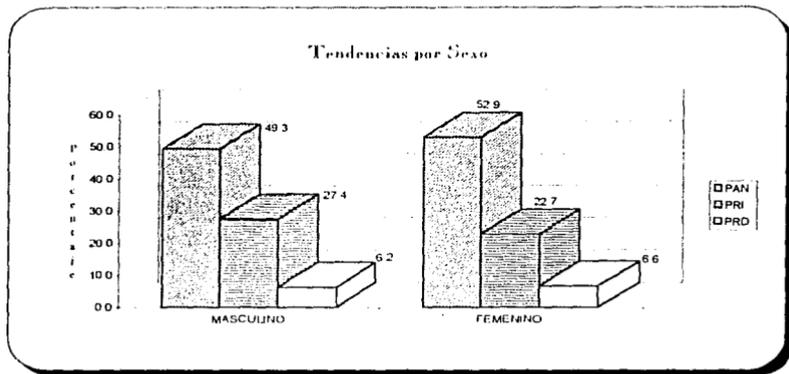


Figura 24. Cruce: Preferencias vs. Sexo: 1261 respuestas acumuladas, 82.6% de 1527 respuestas válidas

Por otra parte, la opción No he Decidido es significativa a tal grado, que supera a los valores obtenidos por el PRD, realmente esta opción provocó que muchas respuestas se desviarán de su verdadero destino (que era algún partido).

Analizando este primer cruce se observa que el PAN goza de mucha popularidad entre los grupos jóvenes, pero se nota que en los grupos: 35 - 44 años y 45 y más la popularidad baja; el PRI tiene altibajos y su valor máximo lo alcanza en el grupo de edad; 45 y más.

En la figura 24 se presentan las preferencias por sexo, una vez más el PAN domina tanto en la clase femenina como en la masculina. Realizando un análisis sobre las 1261 respuestas que registra la figura 24 (recordar que son sólo las acumuladas por PAN, PRI y PRD), 595 son mujeres y 666 hombres, sus porcentajes son 47.1 y 52.9 respectivamente. Por partidos se tiene que de las 595 mujeres, el 64.4% votaría por el PAN, el 27.6% por el PRI y el restante 8% por el PRD; mientras que de los 666 hombres, el 59% se inclina por el PAN, el 33% por el PRI y el 8% por el PRD. En la tabla 21 las proporciones están calculadas sobre un total de 803 hombres y 724 mujeres.

PARTIDO	SEXO	
	HOMBRES	MUJERES
PAN	493	529
PRI	274	227
PRD	62	66
Otro	17	22
Ninguno	25	40
No he Decidido	128	116
Suma Columna	1000	1000
Frecuencias (n)	803.0	724.0

Tabla 21 Cruce Preferencias partidarias vs. Sexo
1527 respuestas

En la figura 25 haciendo un análisis similar al anterior se observa que de las 1198 respuestas que conjuntan los tres partidos en cada nivel de instrucción, 6.8% no tiene estudio alguno, 38.6% estudió hasta la primaria, 23.3% la secundaria, 19% bachillerato, finalmente postbachillerato el 12.3%. Observando las cifras de la figura es curioso cómo va ascendiendo la proporción de la preferencia hacia el PAN hasta el bachillerato y después baja. Para el PRI el comportamiento es prácticamente inverso al del PAN, aquí la proporción de preferencia decrece conforme las personas son más instruidas pero, del bachillerato a postbachillerato hay un ligero incremento que rompe con la tendencia a la baja. La tabla 22 muestra las proporciones de opciones partidarias que no se ven en la figura 25.

Las proporciones correspondientes a los 1170 registros representados en la figura 26 son los siguientes: Desempleados o Jubilados 3%, Ama de Casa 24.4%, Estudiante 4.5%, Empleado 21%, Obrero 8%, Comerciante 18.6%, Ejecutivo 8.3%, Agropecuario 5.7% y Otro 6.4%. El PAN sobresale en los tres grupos de más peso de esta gráfica pero, en Ama de casa (que es el mayor) curiosamente tiene el valor mínimo.

En la tabla 23 se dan las proporciones completas por ocupación. A pesar de todos los esfuerzos para no obtener proporciones bajas existen celdas cuyos valores son insignificantes; por otra parte, hay valores de la opción No he Decidido que superan a los del PRD e incluso a los del PRI.

El último análisis corresponde a las preferencias observadas a través de los ingresos. Estas son las proporciones por rango de ingreso de las 1087 respuestas de la figura 27: 0 a 400 pesos 31.7%, 450 a 1200 pesos 42.1%, 1250 a 2000 pesos 13.9%, 2100 a 3200 pesos 5.4%, 3250 a 4000 pesos 2.4%, y más de 4000 pesos 4.5%. Hay un comportamiento extraño en este cruce porque el rango Más de 4000 pesos tiene mayor número de respuestas que el rango que lo antecede, esto es raro porque lo lógico sería que a mayor intervalo de ingreso menor número de personas, es posible que se deba a un descuido de codificación. En los rangos de mayor respuesta el PAN dominó y obtuvo exactamente la misma proporción en dos de ellos; el PRI estuvo muy

constante en estos mismos rangos; en los grupos restantes se presentaron situaciones extremas, en uno de estos extremos el PAN fue muy superior y, en el otro, el PRI quedó muy atrás; pero se debe tener cuidado al interpretar porque el número de personas para cada grupo no fue mayor a las 70 personas. El PAN estuvo cerca de el 70%, mientras que el PRI no pasó del 20%, y además fue superado por el PRD en uno de los rangos; más información se da en la tabla 24.

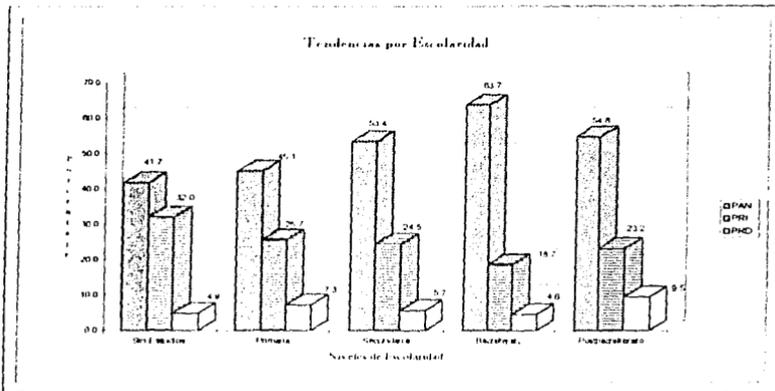


Figura 25. Cruce: Preferencias vs. Escolaridad. 1198 respuestas acumuladas. 82% de 1460 respuestas válidas

PARTIDO	NIVEL DE ESCOLARIDAD				
	Sin Estudios	Primaria	Secundaria	Bachillerato	Post Bachillerato
PAN	41.7	45.1	53.4	63.7	54.8
PRI	32.0	25.7	24.5	18.7	23.2
PRD	4.9	7.3	5.7	4.6	9.5
Otro	0.0	2.0	2.7	1.9	1.2
Ninguno	2.9	3.2	4.5	2.3	2.4
No he Decidido	18.4	16.7	9.3	8.8	8.9
Suma	100.0	100.0	100.0	100.0	100.0
Frecuencias (n)	103.0	592.0	335.0	262.0	168.0

Tabla 22. Cruce: Preferencias partidarias vs. Escolaridad. 1460 respuestas

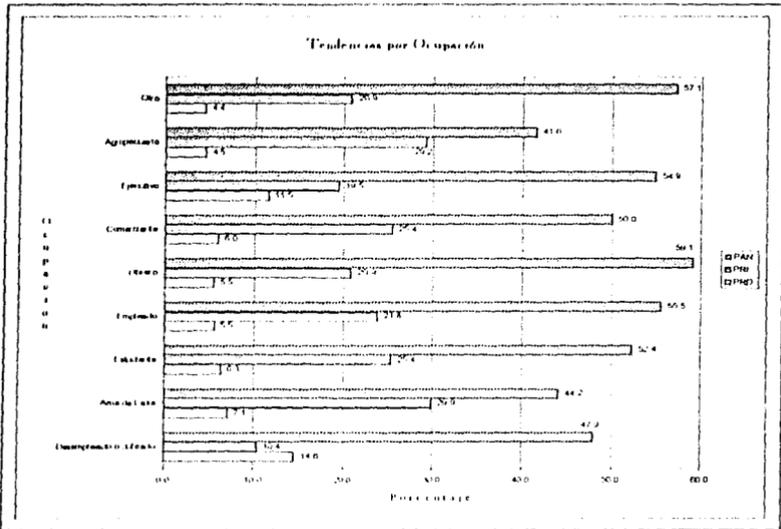


Figura 26. Cruce: Preferencias vs. Ocupación. 1170 respuestas acumuladas, 82.2% de 1423

PARTIDO	O C U P A C I O N E S								
	Desempleado o Jubilado	Ama de Casa	Estudiante	Empleado	Obrero	Comerciante	Ejecutivo	Agropecuario	Otro
PAN	47.9	44.2	52.4	55.5	59.1	50.0	54.9	41.6	57.1
PRD	10.4	29.9	25.4	23.8	20.9	25.4	19.5	29.2	20.9
PRD	14.6	7.1	6.3	5.5	5.5	6.0	11.5	4.5	4.4
Otro	6.3	2.0	3.2	1.4	0.9	2.6	2.7	0.0	1.1
Ninguno	6.3	1.7	1.6	3.8	1.8	3.4	3.5	3.4	5.5
No he decidido	14.6	15.1	11.1	10.0	11.8	12.7	8.0	21.3	11.0
Suma	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Frecuencias (n)	48.0	351.0	63.0	290.0	110.0	268.0	113.0	89.0	91.0

Tabla 23. Cruce: Preferencias partidarias vs. Ocupación. 1423 respuestas

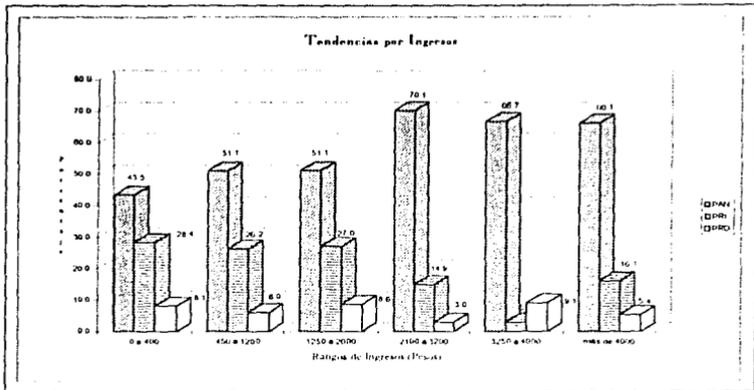


Figura 27. Cruce: Preferencias vs. Ingresos; 1087 respuestas acumuladas. 83% de 1310 respuestas válidas

PARTIDO	RANGOS DE INGRESOS \$ PESOS					
	0 a 400	450 a 1200	1250 a 2000	2100 a 3200	3250 a 4000	más de 4000
PAN	43.5	51.1	51.1	70.1	66.7	66.1
PRD	28.4	26.2	27.0	14.9	3.0	16.1
PRP	8.1	6.0	8.6	3.0	9.1	5.4
Otro	2.1	2.0	1.7	1.5	6.1	1.8
Ninguno	4.2	1.8	4.0	4.5	6.1	1.8
No he Decidido	13.7	12.9	7.5	6.0	9.1	8.9
Suma	100.0	100.0	99.9	100.0	100.1	100.1
Frecuencias (n)	430.0	550.0	174.0	67.0	33.0	56.0

Tabla 24. Cruce: Preferencias partidarias vs. Ingresos 1310 respuestas

Con esto concluye esta subsección y también llegamos al final de la sección, después de analizar todos los cruces y de hacer comentarios particulares, en forma global podemos decir que fueron muy claras las tendencias para cada partido y que los habitantes estaban ansiosos por cumplir con su derecho a votar, quizá esta ansiedad fue provocada por el sentimiento de reafirmar el cambio que deseaban en las elecciones ordinarias y que no sucedió, obligando a realizar un segundo ejercicio electoral.

ANÁLISIS 2

Antes de hacer cualquier interpretación sobre las cifras de las tablas 20A y 21A hay que recordar que: tanto la edad como el sexo fueron variables cuota, y las proporciones de ambas se eligieron de acuerdo a las proporciones globales del estado, y que la cuota del grupo 45 y más años no fue posible satisfacerla (dichas proporciones se encuentran en las figuras 1 y 2 de este capítulo, p. 43).

La tabla 20A muestra distintas distribuciones de proporciones en cuanto a los simpatizantes de cada partido: mientras que el PRD guarda cierta uniformidad en los cuatro grupos de edad, el PAN la tiene por pares de grupos, por su parte el PRI muestra semejanza en los grupos extremos. Finalmente, la opción No he Decidido fue la que mantuvo la "mejor" uniformidad.

PARTIDO	GRUPOS DE EDAD (AÑOS)					Suma	Frecuencias (n)
	18 - 24	25 - 34	35 - 44	45 y más			
PAN	32.9	29.5	18.9	18.7	100.0	774.0	
PRI	26.6	29.4	18.0	26.0	100.0	384.0	
PRD	22.2	25.2	26.3	26.3	100.0	99.0	
PT	17.6	47.1	11.8	23.5	100.0	17.0	
PFCRN	40.0	20.0	0.0	40.0	100.0	5.0	
OTRO	12.5	50.0	12.5	25.0	100.0	8.0	
Ninguno	24.5	24.5	22.4	28.6	100.0	49.0	
No he Decidido	24.1	24.1	25.1	26.7	100.0	187.0	

Tabla 20A. Cruce Preferencias partidarias vs. Grupos de edad. 1523 respuestas

En la tabla 21A se aprecian dentro del PAN y el PRD una distribución muy pareja entre ambos sexos, el PRI muestra una inclinación hacia los hombres y en las opciones no partidistas son más los hombres indecisos y más las mujeres que no piensan sufragar.

Por nivel de escolaridad (tabla 22A) las personas con primaria sobresalieron en todos los partidos. Dentro del PAN los respondientes que estudiaron hasta la secundaria y bachillerato están repartidos similarmente. El PRI muestra un comportamiento igual con respecto a la tabla 22; es decir, mientras más alto es el grado de escolaridad el número de sus simpatizantes baja. Lo notable en el PRD es que tiene una proporción alta en el nivel de mayor escolaridad con respecto a los otros partidos.

PARTIDO	SEXO		Suma	Frecuencias (n)
	HOMBRES	MUJERES		
PAN	50.8	49.2	100.0	779.0
PRI	57.3	42.7	100.0	384.0
PRD	51.0	49.0	100.0	98.0
Otro	46.7	53.3	100.0	30.0
Ninguno	40.8	59.2	100.0	49.0
No he Decidido	55.1	44.9	100.0	187.0

Tabla 21A. Cruce Preferencias partidarias vs. Sexo. 1527 respuestas

PARTIDO	NIVEL DE ESCOLARIDAD						Suma	Frecuencias (n)
	Sin Estudios	Primaria	Secundaria	Bachillerato	Post Bachillerato			
PAN	5.7	35.7	23.9	22.3	12.3	100.0	748.0	
PRI	9.3	42.8	23.1	13.8	11.0	100.0	355.0	
PRD	5.3	45.3	20.0	12.6	16.8	100.0	95.0	
Otro	0.0	42.9	32.1	17.9	7.1	100.0	28.0	
Ninguno	6.4	40.4	31.9	12.8	8.5	100.0	47.0	
No he Decidido	10.2	52.9	16.6	12.3	8.0	100.0	187.0	

Tabla 22A. Preferencias partidarias vs. Escolaridad. 1460 respuestas

La tabla 23A indica que PRD y PRI son partidos de Ama de Casa, el PAN tiene mayor fuerza entre el Empleado, aunque el PRI y el PRD también tiene apoyo en esa categoría, por último, otra clase que sobresale es la de Comerciante, que en promedio acumula junto con las citadas anteriormente el 63.5% de los tres partidos. (Por cuestiones de espacio la estructura de la tabla 23A esta hecha por columnas, pero su propósito no se altera porque los partidos son las columnas.

Los rangos de ingresos con mayor peso son el de 0 a \$400 y el de \$450 a \$1200. El PAN y el PRI alcanzan la proporción mas grande en este último, en cambio el PRD tiene su proporción mayor en el primer rango. De los tres últimos rangos no hay mucho que decir porque sus frecuencias son poco significativas. Las cifras completas están en la tabla 24A.

Así como en la subsección 3.2.3A se convirtieron los salarios mínimos (S.M.) que manejó el INEGI a pesos de 1995, aquí se hace la conversión inversa, es decir, los rangos que usó la encuesta preelectoral se convierten a S.M. mensual vigente en mayo de 1995 de \$556.32.

Los rangos de ingresos de la encuesta preelectoral convertidos según lo antes expuesto quedan: de 0 a menos de un *S.M.*; de menos de un *S.M.* a más de dos *S.M.*; de más de dos *S.M.* a menos de seis *S.M.*; de menos de seis *S.M.* a menos de ocho *S.M.* y el último sería de ocho *S.M.* en adelante al mes.

OCUPACION	PARTIDOS					
	PAN	PRI	PRD	Otro	Ninguno	No he decidido
Desempleado o Jubilado	3.2	1.4	7.4	10.7	6.8	3.9
Ama de Casa	21.5	29.7	26.3	25.0	13.6	29.3
Estudiante	4.6	4.5	4.2	7.1	2.3	3.9
Empleado	22.3	19.5	16.8	14.3	25.0	16.0
Obrero	9.0	6.5	6.3	3.6	4.5	7.2
Comerciante	18.5	19.3	16.8	25.0	20.5	18.8
Ejecutivo	8.6	6.2	13.7	10.7	9.1	5.0
Agropecuario	5.1	7.4	4.2	0.0	6.8	10.4
Otro	7.2	5.4	4.2	3.6	11.4	5.5
Suma	100.0	99.9	99.9	100.0	100.0	100.0
Frecuencias (n)	722.0	353.0	95.0	28.0	44.0	181.0

Tabla 23A. Cruce Preferencias partidarias vs. Ocupación. 1423 respuestas

PARTIDO	RANGOS DE INGRESOS S PESOS						Suma	Frecuencias (n)
	0 a 400	450 a 1200	1250 a 2000	2100 a 3200	3250 a 4000	más de 4000		
PAN	28.2	42.4	13.4	7.1	3.3	5.6	100.0	663.0
PRI	36.6	43.2	14.1	3.0	0.3	2.7	100.0	333.0
PRD	38.5	36.3	16.5	2.2	3.3	3.3	100.0	91.0
Otro	33.3	40.7	11.1	3.7	7.4	3.7	100.0	27.0
Ninguno	43.9	24.4	17.1	7.3	4.9	2.4	100.0	41.0
No he Decidido	38.1	45.8	8.4	2.6	1.9	3.2	100.0	155.0

Tabla 24A. Cruce Preferencias partidarias vs. Ingresos. 1310 respuestas

3.2.4 Preferencias por Partido en cada Estrato

En esta sección se examinan las tendencias dentro del marco de la estratificación; para tal efecto se dividió el estado en siete estratos, se consideran las mismas opciones partidarias que en la subsección 3.2.3B; el total poblacional para este análisis es de 1541 respuestas válidas.

Conviene recordar la estratificación manejada: estrato 1, con 12 municipios cuyas poblaciones son de hasta 10,000 ciudadanos (se muestrearon 6 municipios); estrato 2, con 16 municipios con poblaciones desde 10,001 hasta 30,000 ciudadanos (se muestrearon 12); estrato 3, con 14 municipios con poblaciones desde 30,001 hasta 70,000 ciudadanos, (se seleccionaron todos); estrato 4, Salamanca con 109,438 ciudadanos; estrato 5, Celaya con 163,012 ciudadanos; estrato 6, Irapuato con 187,427 ciudadanos; y estrato 7, León con 446,299 ciudadanos.

En las tablas 25 y 26 se encuentran las tendencias de la encuesta y los resultados oficiales respectivamente; los valores se presentan en proporciones para cada estrato; en la tabla de la encuesta están las proporciones de las seis opciones posibles sin considerar el diseño muestral, y en la de los resultados se fusionan el PT y el PFCRN bajo el concepto OTRO.

En la tabla 25 se aprecia cómo hay saltos grandes del estrato uno al dos, sobretudo para el PAN, PRI y PRD. Por partidos, el PAN muestra decrecimiento del estrato dos al cinco y después repunta; el PRI baja consistentemente desde el estrato uno hasta el cuatro y en el cinco sube en forma un tanto brusca de acuerdo a su comportamiento, por último, baja en Irapuato y se incrementa casi en un punto en León. El PRD tiene un comportamiento muy irregular, porque de no tener nada en el estrato uno tiene un alza sostenida hasta alcanzar su máximo en Salamanca y después vuelve a sufrir altibajos; la respuesta Sin Decidir merma a los partidos porque alcanza valores significativos. Para terminar, hay que notar que la diferencia entre el PAN y el PRI en el estrato uno no es tan grande como generalmente se venía dando en análisis anteriores.

ENCUESTA PREELECTORAL TENDENCIAS POR ESTRATOS						
ESTRATO	PAN	PRI	PRD	OTRO	NINGUNO	SIN DECIDIR
1	41.5	33.8	0.0	1.5	6.2	16.9
2	49.4	25.3	6.4	0.8	3.0	15.1
3	49.2	24.4	9.2	1.7	2.7	12.8
4	48.8	21.3	11.3	1.3	2.5	15.0
5	46.7	27.0	4.1	3.3	4.1	14.8
6	53.0	23.9	6.0	3.0	6.0	8.2
7	56.7	24.8	3.7	2.6	2.6	9.7
TOTAL	50.75	25.05	6.49	1.95	3.24	12.52

Tabla 25. Tendencias partidarias por estratos registradas por la encuesta preelectoral; 1541 respuestas en total sin considerar diseño de muestreo

La tabla 26 muestra algo sorprendente, el PRI ganó en el estrato uno; independientemente de que son resultados oficiales, esto es interesante porque se están considerando los doce municipios que lo forman; ahora bien, aunque no se trata de una diferencia grande y

considerando que el peso del estrato uno según se ve en la tabla 26 es del 3.4%, si es llamativa por el claro dominio que ha tenido el PAN en la presente encuesta.

RESULTADOS OFICIALES TENDENCIAS POR ESTRATOS*					Peso del Estrato
	PAN	PRI	PRD	OTRO	
ESTRATO 1	45.2	48.2	5.0	1.7	3.4
ESTRATO 2	54.0	37.1	7.1	1.8	16.6
ESTRATO 3	51.9	36.4	9.6	2.1	33.9
ESTRATO 4	54.9	32.4	10.7	2.0	5.7
ESTRATO 5	61.8	30.6	4.5	3.1	8.0
ESTRATO 6	62.3	28.0	7.6	2.1	9.4
ESTRATO 7	67.5	27.0	3.9	1.6	23.0
TOTAL	58.10	32.90	7.02	1.98	100.0

* Se consideran todos los municipios que conforman cada estrato

Tabla 26. Resultados oficiales en los estratos diseñados para la preelectoral: 1,244,955 votos válidos. IEE Guanajuato

Las tablas 27 y 28 muestran únicamente las proporciones sobre el acumulado obtenido por partidos. La tabla 27 no está hecha bajo el diseño de muestreo. La tabla 28 no considera la totalidad de los municipios que forman cada estrato, sino que está trasladada a los municipios que salieron en muestra en cada estrato para efectos de la encuesta preelectoral.

ENCUESTA PREELECTORAL TENDENCIAS POR ESTRATOS (SOLO PARTIDOS POLÍTICOS)				
	PAN	PRI	PRD	OTRO
ESTRATO 1	54.0	44.0	0.0	2.0
ESTRATO 2	60.4	30.9	7.8	0.9
ESTRATO 3	58.2	28.9	10.8	2.0
ESTRATO 4	59.1	25.8	13.6	1.5
ESTRATO 5	57.6	33.3	5.1	4.0
ESTRATO 6	61.7	27.8	7.0	3.5
ESTRATO 7	64.6	28.2	4.2	2.9
TOTAL	60.25	29.74	7.70	2.31

Tabla 27. Tendencias partidarias por estratos, exclusivamente partidos políticos: 1298 respuestas, 84.23% de las 1541 totales sin diseño de muestreo

La tabla 28 muestra en los estratos 1 y 2 diferencias muy parecidas entre PAN y PRI. Recordando que el PRI ganó en el estrato uno (tabla 26), es necesario dar más atención a lo que sucede entre estos dos partidos en el mismo estrato, tanto en la encuesta de salida como en el conteo rápido. Por último, se observa cómo mientras el PRI tiene una baja consistente conforme cambia de estrato, el PAN sólo tiene una baja del estrato uno al dos, y después crece paulatinamente.

En la figura 28 se ilustra la comparación numérica de los municipios muestreados de cada estrato entre la encuesta preelectoral (sin diseño muestral) y los resultados oficiales.

RESULTADOS OFICIALES TENDENCIAS POR ESTRATOS**				
	PAN	PRI	PRD	OTRO
ESTRATO 1	52.4	39.8	6.3	1.5
ESTRATO 2	51.3	38.5	8.2	2.1
ESTRATO 3	51.9	36.4	9.6	2.1
ESTRATO 4	54.9	32.4	10.7	2.0
ESTRATO 5	61.8	30.6	4.5	3.1
ESTRATO 6	62.3	28.0	7.6	2.1
ESTRATO 7	67.5	27.0	3.9	1.6
TOTAL	58.43	32.36	7.19	2.02
** Se consideran los mismos municipios seleccionados en cada estrato por la encuesta preelectoral				

Tabla 28. Resultados oficiales en los municipios muestreados por estrato; 1,153,991 votos válidos. IEE Guanajuato

La tabla 29 contiene los resultados obtenidos con el diseño muestral de las proporciones para cada partido, sus varianzas e intervalos de confianza, en relación a estos últimos, se calcularon sustituyendo los valores de las varianzas que se encuentran en el mismo renglón de cada intervalo en la fórmula 3.19 de la sección 3.1.3. El muestreo estratificado bietápico se aplicó también sobre 1541 respuestas y sólo presenta lo concerniente a los partidos políticos.

Finalmente, la tabla 30 muestra el comparativo entre los valores totales obtenidos mediante el muestreo estratificado bietápico (tabla 29) y los totales sin considerarlo. A simple vista se observa poca diferencia entre uno y otro.

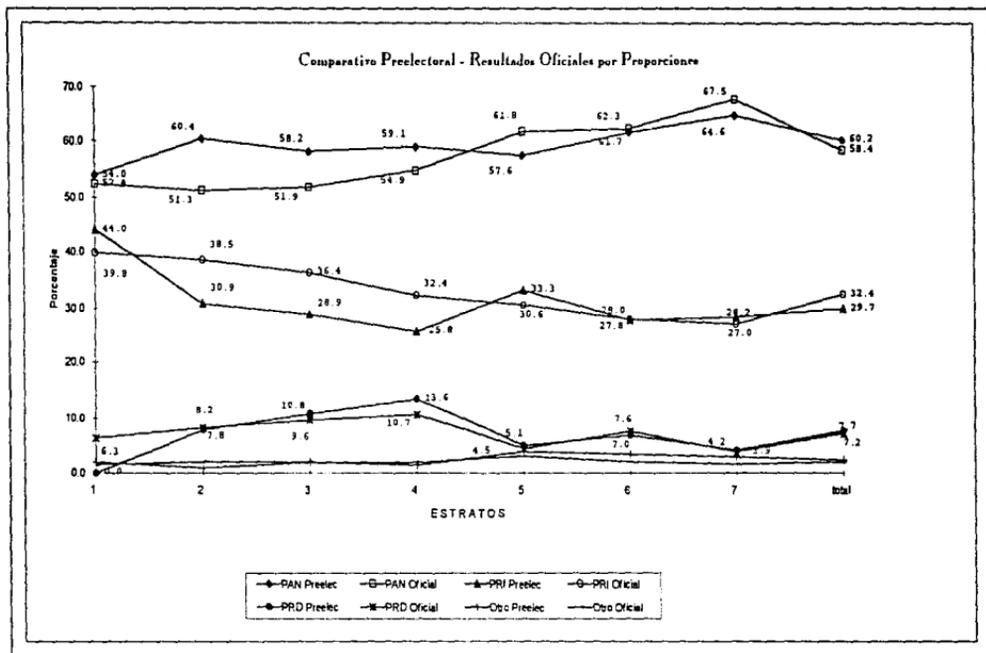


Figura 28. Comparación entre resultados oficiales y encuesta preelectoral por estratos basada en las tablas 27 y 28

ENCUESTA PREELECTORAL			
TENDENCIAS CON MUESTREO ESTRATIFICADO BIETÁPICO			
PARTIDO	ESTIMADORES		INTERVALOS DE CONFIANZA
	MEDIA POBLACIONAL, \bar{Y}_w	VARIANZA $\hat{var}(\bar{Y}_w)$	Al 95%
PAN	50.6	2.00×10^{-4}	(47.83, 53.38)
PRI	24.9	6.20×10^{-5}	(23.34, 26.45)
PRD	6.6	7.25×10^{-6}	(6.12, 7.17)
Otro	2.0	5.65×10^{-7}	(1.87, 2.16)

Tabla 29. Tendencias partidarias considerando el diseño de muestreo (Estratificado Bietaipico descrito en la sección 3.1.1)

GUANAJUATO			
COMPARATIVO GLOBAL TENDENCIAS - OFICIALES			
ENCUESTA PREELECTORAL (Mayo 1995)			
PARTIDO	SIN DISEÑO	CON DISEÑO	OFICIAL
PAN	50.75	50.60	58.10
PRI	25.05	24.91	32.90
PRD	6.49	6.64	7.02
Otro	1.95	2.02	1.98

Tabla 30. Resultados globales de encuesta preelectoral, (sin y considerando diseño de muestreo) y resultados oficiales (IEE Guanajuato, Elección extraordinaria para gobernador 1995)

La expresión gráfica de la tabla 30 se presenta en la figura 29, aquí se aprecia la importancia que adquiere el momento en el que se aplica una encuesta preelectoral. Aunque faltaban dos semanas para la celebración de las elecciones, los resultados finales de la encuesta preelectoral distaban de las proporciones reales que alcanzaron el PAN y el PRI, aun así dejaban claro un escalafón. En cuanto al PRD y los otros partidos sería erróneo interpretar los valores obtenidos como una excelente precisión de la estimación. La cercanía numérica entre los resultados oficiales y los de la encuesta se debe a la baja fuerza de atracción que ejercieron el PRD y los otros partidos sobre el electorado del estado.

En términos generales la encuesta logra crear un precedente en relación a las preferencias para cada partido político y por ende el acomodamiento de cada uno en cierto orden.

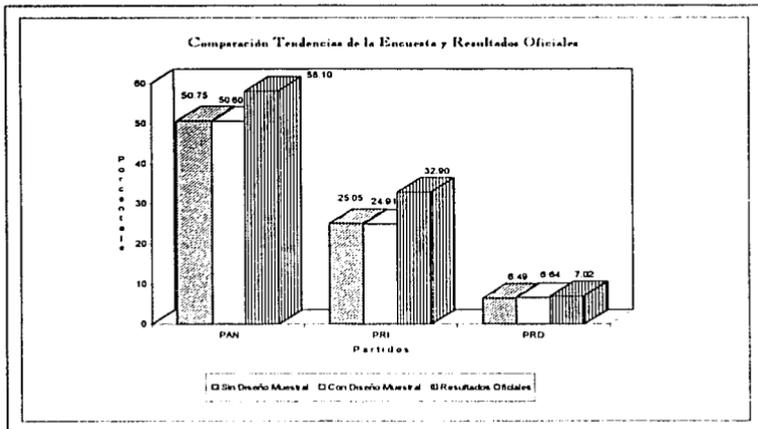


Figura 29. Comparación entre resultados oficiales y encuesta preelectoral considerando y sin considerar diseño muestral

3.2.5 Preferencia Partidista

En esta sección se hará un seguimiento de los tres partidos más importante en cuanto a la popularidad de cada uno. Las siguientes figuras muestran la preferencia para cada partido en forma abreviada, porque las preferencias para sus rivales se engloban en el concepto OPOSICION; también se encuentran las respuestas NINGUNO y SIN DECIDIR con el propósito de ilustrar la totalidad de las respuestas. Los resultados oficiales están manejados de igual forma.

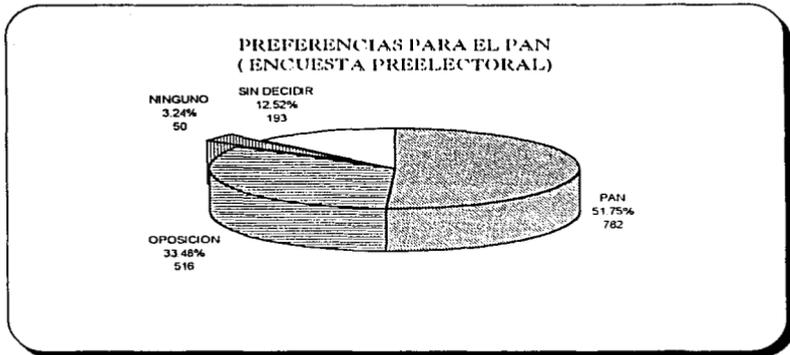


Figura 30. Preferencias Partido Acción Nacional, 1541 respuestas totales

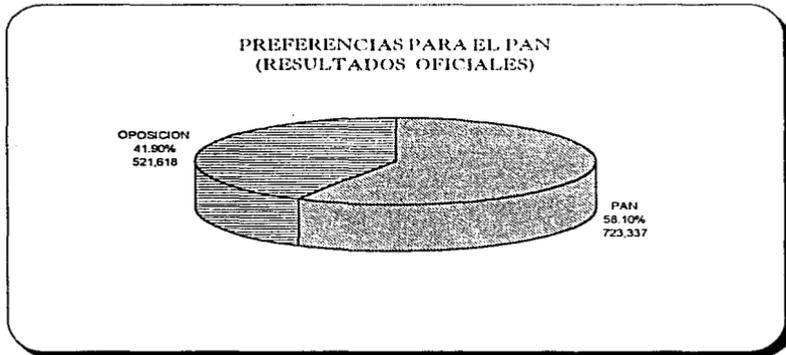


Figura 31. Resultados oficiales; preferencias Partido Acción Nacional, 1,244,955 votos válidos

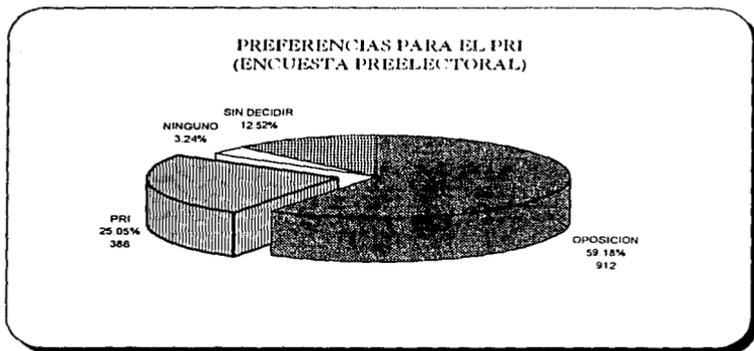


Figura 32. Preferencias Partido Revolucionario Institucional, 1541 respuestas totales

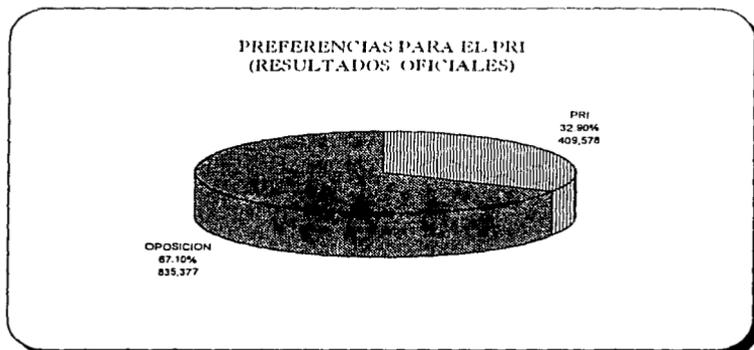


Figura 33. Resultados oficiales: preferencias Partido Revolucionario Institucional, 1,224,955 votos válidos

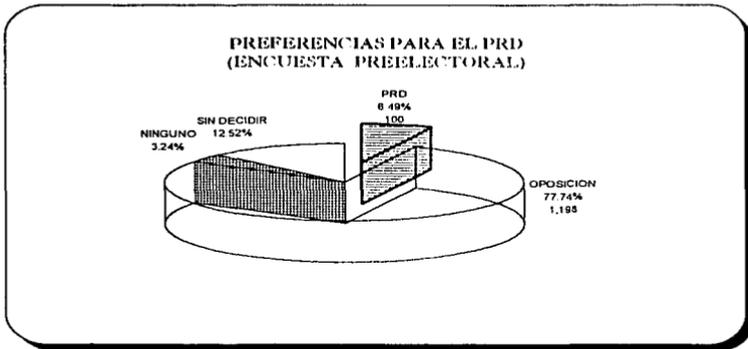


Figura 34. Preferencias Partido de la Revolución Democrática, 1541 respuestas totales

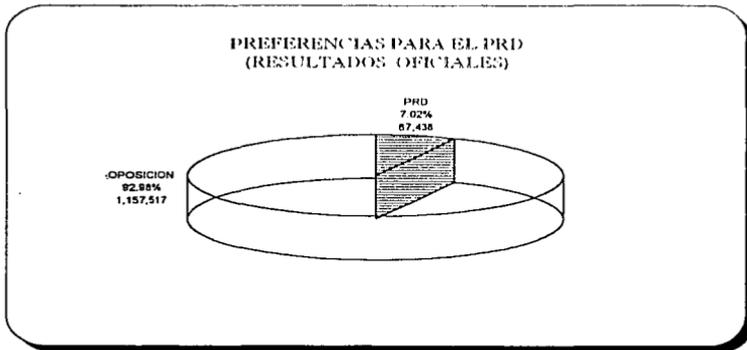


Figura 35. Resultados oficiales: preferencias Partido de la Revolución Democrática, 1,224,955 votos válidos

Ahora se presenta una pregunta de la encuesta que se puede considerar como indicador de la popularidad y por ende de la preferencia, la pregunta fue ¿Conoce al candidato del partido "X"? Los resultados están en las siguientes gráficas. Todas las figuras están calculadas sobre 1541 respuestas.

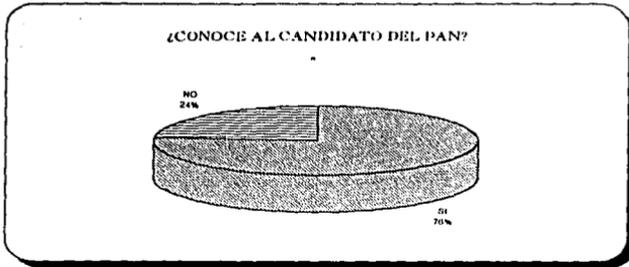


Figura 36A

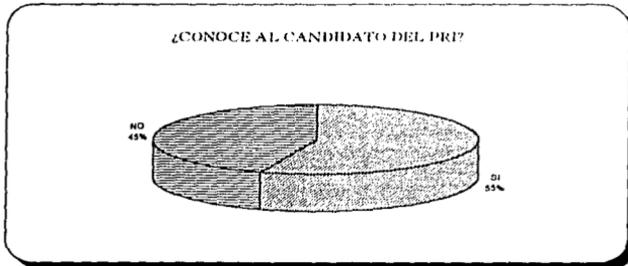


Figura 36B

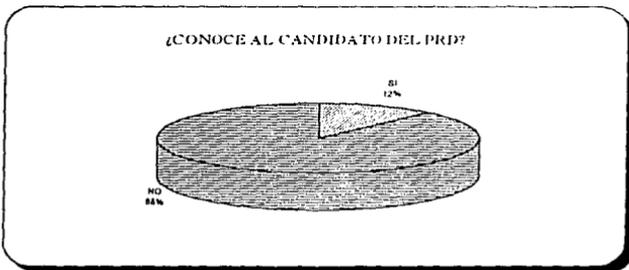


Figura 36C

Otra pregunta interesante para esta sección fue ¿A quién considera el candidato idóneo para asumir el gobierno del estado?

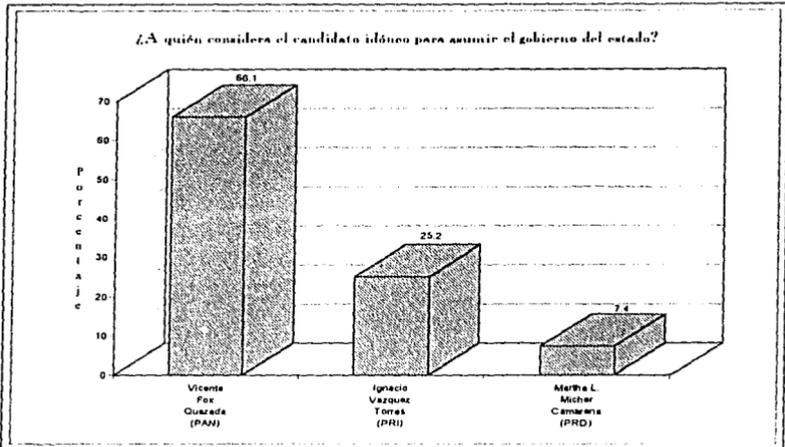


Figura 37. Preferencias para cada candidato, 1283 respuestas acumuladas, 98.7% de 1300 respuestas válidas

Tal vez se considere a esta sección inútil, porque repite lo que se ha venido evidenciando desde la subsección 3.2.3A pero, realmente tiene una óptica diferente e incluso resumida; se puede utilizar para saber rápidamente la simpatía de los partidos y de sus respectivos candidatos en la población del estado, ya que en la encuesta no se incluyó pregunta alguna para conocer la opinión de la ciudadanía acerca de si su voto lo emitirán por el candidato o por el partido.

Considerando 1604 respuestas las proporciones de la figura 37 son: para el candidato panista, 53.6%; para el priista, 20.4%; para la perredista, 6.0%; para Salvador Arévalo Maldonado del PT, 0.8% y para Israel González Arreguín candidato del PFCRN el 0.2%; por último, no contestaron el 19.0% de los entrevistados.

En otro orden de ideas se hizo el cruce entre la pregunta ¿Por qué partido votaría? y ¿Conoce al candidato del partido "X"?, los resultados son los siguientes: de 788 personas que dijeron votarían por el PAN, 710 contestaron que sí conocían a su candidato, esto es el 90%; de 387 simpatizantes del PRI, 250 sí conocen a su candidato, casi el 65%; para el PRD tenemos 100 adeptos, de los cuales 40 conocían a su candidato. Para finalizar, cuando a las 193 personas que estaban indecisas se les preguntó si conocían al candidato del PAN, el 67% dijo que sí y, de éstas mismas, el 55% declararon conocer al candidato priista.

Sólo resta mencionar que el PAN tenía supremacía sobre la mayoría de la población guanajuatense el día en que se levantó la encuesta preelectoral, a pesar de que el PRI demostró una ventaja en uno de los estratos (tabla 26 de este capítulo).

3.2.6 Repulsión Partidista

Esta sección es el complemento de la anterior, aquí se analizará el rechazo a los partidos, primero se hará en forma global y después por estratos. La tabla 31 contiene las frecuencias válidas que, para efectos de la tabla, se llamarán *Antivotos* y las proporciones correspondientes.

ENCUESTA PREELECTORAL REPULSIÓN PARTIDISTA		
PARTIDO	Partido por el que no votaría	
	ANTIVOTOS	PROPORCION
PAN	169	10.9
PRI	713	45.8
PRD	270	17.3
OTRO	206	13.2
NINGUNO	199	12.8
TOTAL	1,557	100.0

Tabla 31. Rechazo partidista, encuesta preelectoral

La tabla anterior muestra que el PRI tiene un alto grado de rechazo, mientras que el PAN registra el más bajo.

ENCUESTA PREELECTORAL REPULSIÓN PARTIDISTA POR ESTRATOS									
ESTRATO	Partido por el que no votaría								Prop. por estrato
	PAN	PRI	PRD	PT	PFCRN	OTRO	Ninguno		
1	10.9	28.1	15.6	10.9	14.1	1.6	18.8	4.1	
2	11.3	47.2	18.1	4.2	2.3	2.3	14.7	17.1	
3	10.5	42.1	20.5	6.2	5.1	0.9	14.7	34.3	
4	7.4	55.6	14.8	2.5	6.2	4.9	8.6	5.2	
5	9.8	48.0	17.9	3.3	11.4	4.1	5.7	7.9	
6	9.0	51.5	17.9	6.7	4.5	3.0	7.5	8.7	
7	13.1	48.0	12.6	6.0	5.1	2.3	12.9	22.6	
Prop. por Partido	10.9	45.7	17.4	5.6	5.5	2.1	12.8	100.0	

Tabla 32. Rechazo partidista en cada estrato, 1549 respuestas válidas

La tabla 32 desglosa la repulsión partidista por estratos; en la columna de la extrema derecha se encuentran las proporciones de cada estrato sobre el total de la muestra y el último renglón representa lo mismo pero por partido. A continuación se presentaran las figuras por estrato derivadas de la misma tabla.

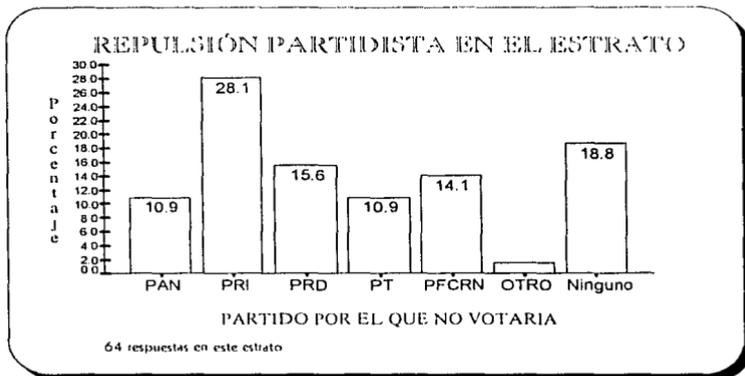


Figura 38A. Rechazo partidista en el estrato 1, representa el 4.1% de la encuesta

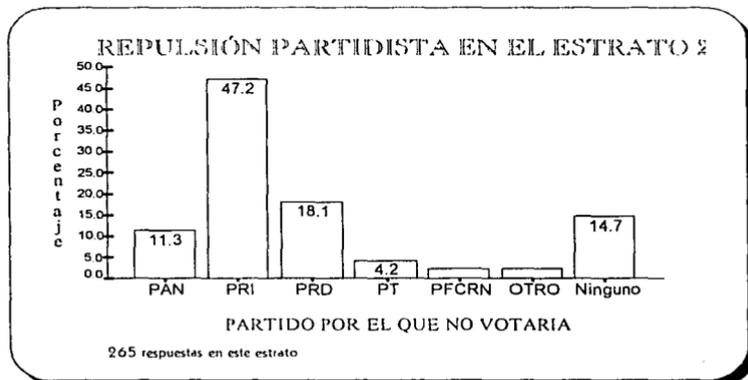


Figura 38B. Rechazo partidista en el estrato 2, representa el 17.1% de la encuesta

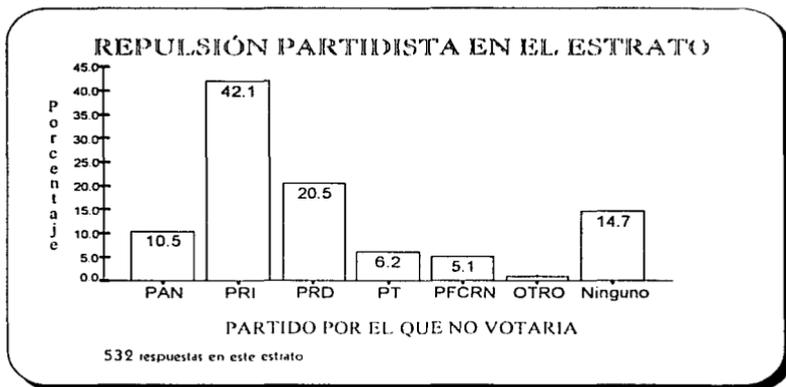


Figura 38C. Rechazo partidista en el estrato 3, representa el 34.3% de la encuesta

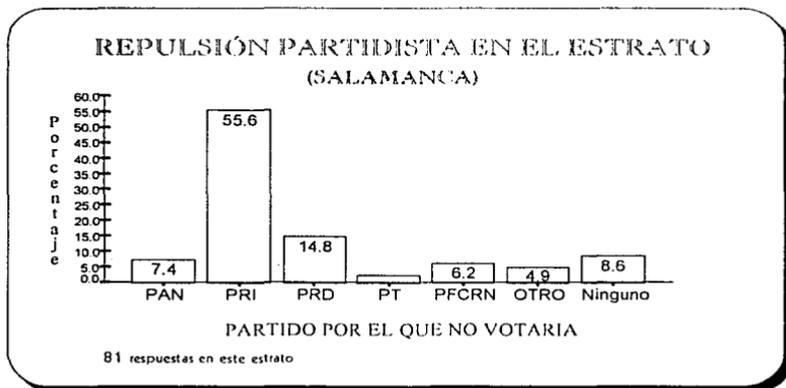


Figura 38D. Rechazo partidista en Salamanca, representa el 5.2% de la encuesta

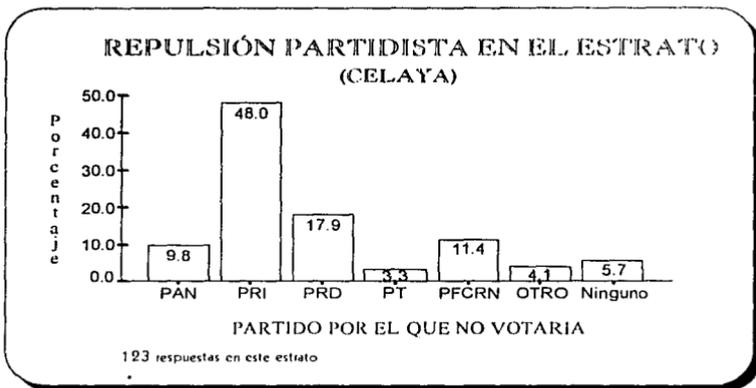


Figura 38E. Rechazo partidista en Celaya, representa el 7.9% de la encuesta

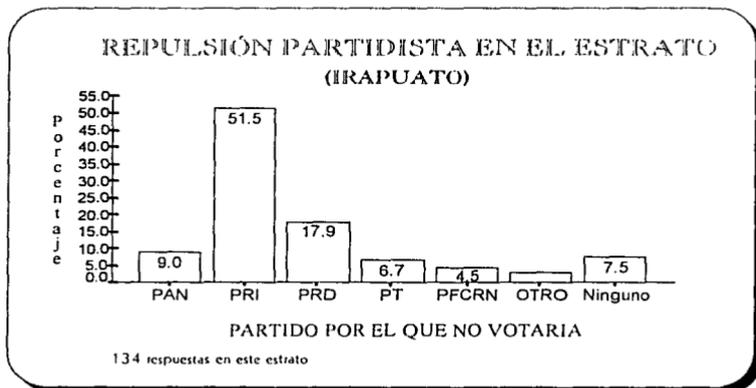


Figura 38F. Rechazo partidista en Irapuato, representa el 8.7% de la encuesta



Figura 38G. Rechazo partidista en León, representa el 22.6% de la encuesta

Como era de esperarse y siguiendo el mismo patrón de repudio que registró el PRI a nivel general, en los estratos sucede algo similar, tal es el caso de Salamanca donde el grado de rechazo es muy alto, en todos los demás con excepción del estrato 1, el PRI tiene en promedio el 47% de desprecio, esto se convierte en otro factor favorable para el PAN y, tal vez, fue algo que no aprovechó el PRD, incluso para lograr colocarse como la segunda fuerza política en el estado.

3.2.7 Comparativo Preferencia vs. Rechazo

En esta sección se utilizan los datos de las dos anteriores para procesarlos de una forma distinta y tratar de establecer si existe repulsión entre los simpatizantes de los partidos. Para tales efectos la tabla 33 muestra el cruce entre las preguntas ¿Partido por el que votaría? y ¿Partido por el que no votaría? Esta tabla esta construida en base a 1477 respuestas válidas que son el 92% de la encuesta.

La forma de interpretar la tabla es la siguiente: se toman las respuestas a la pregunta ¿Partido por el que votaría?, como simpatizantes de la opción que hayan seleccionado; al mismo tiempo, a los que respondieron la pregunta ¿Partido por el que no votaría?, se les considera como repulsivos a la opción escogida. Esto permite acomodar a los simpatizantes de un solo partido por renglón y observar por cuál partido sienten mayor repudio. Cabe señalar que las celdas que contienen el cruce del mismo partido están coloreadas o bloqueadas y que no se consideraron respuestas válidas, tampoco se consideraron válidas aquellas respuestas que se perdieron al fusionar PT y PFCRN con OTRO (aunque en una tabla construida con las opciones originales, si hubieran sido válidas).

ENCUESTA PREELECTORAL REPULSIÓN ENTRE PARTIDOS					
PARTIDO POR EL QUE VOTARIA	PARTIDO POR EL QUE NO VOTARIA				
	PAN	PRI	PRD	OTRO	Ninguno
PAN		70.1	14.6	10.7	4.6
PRI	30.0		33.5	20.2	16.3
PRD	12.4	70.1		6.2	11.3
OTRO	19.2	69.2	7.7		3.8
Ninguno	4.2	33.3	14.6	6.3	
No he decidido	6.0	31.0	12.5	16.8	33.7

Tabla 33. Rechazo entre partidos, las proporciones son sobre los simpatizantes de cada partido totalizando 1477 respuestas válidas distribuidas por renglon así: 755 del PAN (51.1%); 367 del PRI (24.8%); 97 del PRD (6.6%); 26 de Otros (1.8%); 48 no van a votar (3.2%) y 184 aun no deciden (12.5%)

Para ampliar la visión de esta sección se muestran cuatro gráficas (figuras 39 a 42) cuya estructura está basada en las opciones originales de las preguntas involucradas en el cruce; las tres primeras (39 - 41) representan a las personas que simpatizan con los tres partidos más importantes, respectivamente. La figura 42, contiene las proporciones de la opción No he

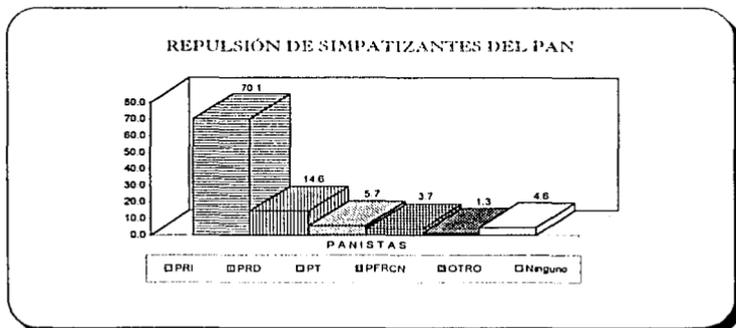


Figura 39. Repulsión del PAN contra los demás. 755 simpatizantes del PAN representan el 51.1% de la encuesta

Decidido; tal vez con la ayuda de ésta se pueda vislumbrar si en efecto estaban indecisos, si ya tenían una clara opción o, definitivamente, no iban a participar en los comicios.

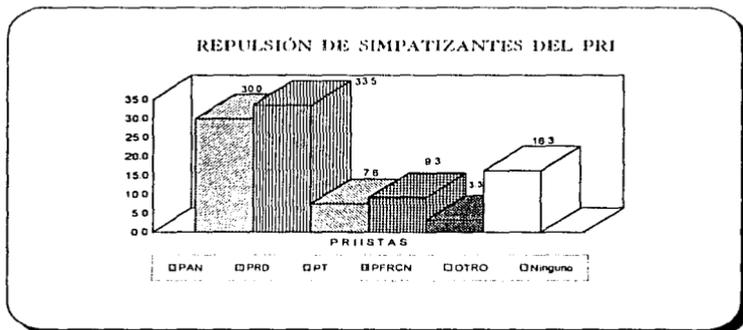


Figura 40. Repulsión del PRI contra los demás. 367 simpatizantes del PRI representan el 24.8% de la encuesta

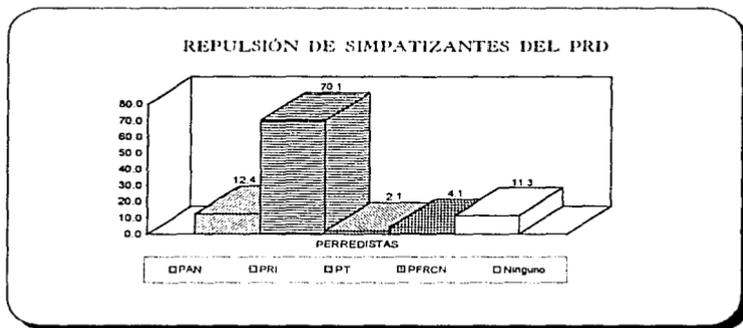


Figura 41. Repulsión del PRD contra los demás. 97 simpatizantes del PRD representan

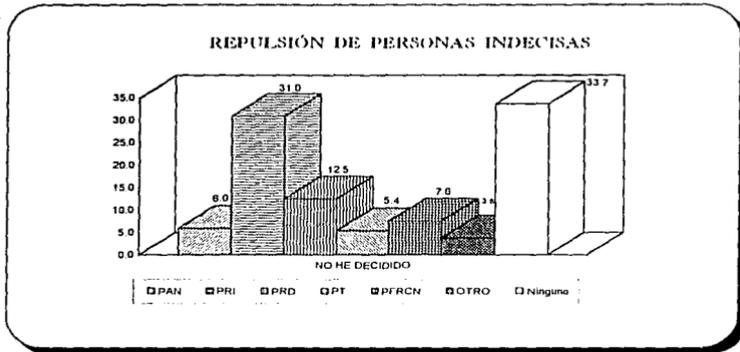


Figura 42. Repulsión de 184 personas que no han decidido, representan el 12.5% de la encuesta

Después de observar las figuras anteriores era seguro que los seguidores de los diversos partidos, y hasta los que no lo eran, estaban más interesados en la derrota del PRI que en la victoria de su propia organización política; es extremadamente notorio el rencor que tenían contra el PRI. Por otro lado es admirable que con tales grados de rechazo se haya mantenido en la segunda posición hasta el final. En otro punto, los priistas consideraban al PRD su acérrimo enemigo mientras que el verdadero rival a vencer era el PAN.

La figura 42 muestra dos aspectos importantes; el primero es que el PRI causaba (por lo menos en ese momento) antipatía a las personas indecisas, el segundo es que en el momento de la aplicación de la encuesta la mayoría tenía un perfil apolítico.

CAPITULO 4 Encuesta de Salida

4.1 Introducción

Este capítulo abarca la encuesta de salida, la estructura es similar a la del capítulo anterior y la diferencia es que aquí no se manejará ningún aspecto sociodemográfico, únicamente se describirá el diseño muestral utilizado y se desglosaran los resultados obtenidos con algunos comentarios adicionales.

4.1.1 Descripción

La encuesta de salida es un proceso que consiste en llenar una cuota con respuestas durante cierto tiempo, en realidad el mecanismo operativo es sencillo, lo único que tiene que hacer el enumerador es esperar a que los votantes salgan después de emitir su voto, y preguntar si es posible saber por cual partido político votó, para así registrar su voto hasta completar la cuota. Algunos problemas comunes que los encuestadores enfrentan son: que en algunas ocasiones no se pueden llenar las cuotas porque los votantes no acceden a responder; al entrevistador se le prohíbe llevar a cabo su trabajo; y en situaciones de alto abstencionismo, no es posible completar la cuota por la escasa asistencia en el período de tiempo asignado.

4.1.2 Diseño Muestral

En cuanto al procedimiento muestral, se aplicó un muestreo de dos etapas con selección sistemática y arranque aleatorio.

A continuación se hace una breve explicación de la selección sistemática con arranque aleatorio: supongamos que se numeran las unidades de la población desde 1 hasta N en cierto orden. Se extrae una muestra que consta de n unidades; se elige una al azar de entre las primeras k ; después las subsecuentes, se toman en múltiplos de k . Por ejemplo, digamos que se escogen de entre las primeras 20 unidades, aleatoriamente sale el número 11, entonces las siguientes serán $11+20$, $11+2(20)$, $11+3(20)$ y así sucesivamente. Es claro que la selección de la primera unidad determina la muestra.

En la práctica, será muy difícil que N sea un múltiplo entero de k , esto provoca que algunas de las muestras sistemáticas no sean del mismo tamaño.

Es necesario hacer algunas consideraciones, el muestreo sistemático facilita la extracción de una muestra con baja posibilidad de cometer errores; también, como implícitamente estratifica, es posible suponer que es más preciso que el muestreo aleatorio simple, porque divide a la población en n estratos conformados por las primeras k unidades, las segundas k , etc.; sin embargo, en el muestreo aleatorio estratificado la posición se determina de una forma aleatoria dentro de cada estrato, mientras que en el muestreo sistemático esta posición es fija, es decir, siempre ocurre en la misma posición relativa del estrato; por último, con el muestreo sistemático se reparte más uniformemente la extracción sobre la población.

Aunque el muestreo aleatorio estratificado y el sistemático son semejantes, no son equivalentes, por lo que, se debe ser cuidadoso al realizar comparaciones entre ambos métodos.

El diseño muestral es relativamente parecido al de la encuesta preelectoral, en el aspecto de que también es bietápico. Para la encuesta de salida, las unidades primarias son los votos válidos de una sección, y las subunidades son una muestra aleatoria de votos válidos de la misma; la diferencia obvia es que no hay estratificación. En la tabla 1 se presentan: el número de secciones seleccionadas, el número que le fue asignado por el IFE a cada sección, el municipio al que pertenece, los votos válidos registrados y muestreados en cada sección y la proporción de votos válidos muestreados sobre el total por sección. La proporcionalidad, aleatoriedad y uniformidad de la distribución muestral se logra automáticamente, gracias a los efectos de la selección sistemática con arranque aleatorio.

ENCUESTA DE SALIDA DISTRIBUCIÓN DE LA MUESTRA					
	Número de Sección	Municipio al que Pertenece	Votos Válidos en cada Sección (M.)	Votos Válidos Muestreados en la Sección (m.)	Proporción sobre total de Votos Válidos de la Sección
n					
1	2945	XICHU	169	30	17.8%
2	2026	PUEBLO NUEVO	623	30	4.8%
3	2736	TARANDACUAO	309	40	12.9%
4	0750	DR. MORA	310	60	19.4%
5	2897	VICTORIA	190	29	15.3%
6	1894	OCAMPO	431	30	7.0%
7	0922	HUANIMARO	603	31	5.1%
8	0725	CUERAMARO	342	43	12.6%
9	2350	S. D. DE LA UNION	431	29	6.7%
10	1183	JARAL DEL PROGR	612	30	4.9%
11	2054	PURISIMA DEL R.	320	30	9.4%
12	0570	CD. M. DOBLADO	487	30	6.2%
13	0575	"	544	30	5.5%
14	2796	URIANGATO	568	30	5.3%
15	2100	ROMITA	161	36	22.4%
16	1204	JRECUARO	331	30	9.1%
17	1208	"	368	14	3.8%
18	1233	"	184	7	3.8%
19	0250	APASEO EL ALTO	364	32	8.8%
20	0254	"	411	30	7.3%
21	1878	MOROLEON	463	21	4.5%
22	0617	COMONFORT	672	30	4.5%
23	0618	"	369	30	8.1%
24	2625	STA. CRUZ DE J. R.	759	30	4.0%
25	0296	APASEO EL GDE.	658	30	4.6%
26	0297	"	545	30	5.5%
27	2951	YURIRIA	423	28	6.6%
28	2957	"	317	31	9.8%
29	0012	ABASOLO	537	10	1.9%
30	0017	"	461	30	6.5%
31	0050	"	470	30	6.4%
32	2545	S. LUIS DE LA PAZ	647	30	4.6%
33	2550	"	205	29	14.1%
34	2553	"	746	30	4.0%
35	2600	"	277	30	10.8%
36	2450	SAN FCO. DEL RIN.	488	26	5.3%
37	2475	"	349	31	8.9%
38	2500	"	212	29	13.7%

Tabla 1. Distribución muestral para la Encuesta de Salida

Tabla 1 (Continuación)

n	Número de la Sección	Municipio al que Pertenece	Votos Válidos en la Sección (M.)	Votos Válidos Muestreados (m.)	Proporción sobre total m_i / M_i
39	2377	SAN FELIPE	424	30	7.1%
40	2383	"	521	30	5.8%
41	2386	"	436	30	6.9%
42	0663	CORTAZAR	502	30	6.0%
43	0675	"	449	30	6.7%
44	2269	SALVATIERRA	458	30	6.6%
45	2270	"	528	30	5.7%
46	2275	"	446	20	4.5%
47	0756	DOLORES HGO.	737	21	2.8%
48	0762	"	765	30	3.9%
49	0763	"	569	27	4.7%
50	0150	ALLENDE	508	29	5.7%
51	0151	"	439	28	6.4%
52	0153	"	709	29	4.1%
53	0162	"	370	28	7.6%
54	0062	ACAMBARO	478	31	6.5%
55	0075	"	322	31	9.6%
56	0125	"	153	32	20.9%
57	2663	SILAO	772	33	4.3%
58	2668	"	665	36	5.4%
59	2712	"	674	30	4.5%
60	2820	VALLE DE SNTGO.	501	24	4.8%
61	2845	"	377	30	8.0%
62	2870	"	530	30	5.7%
63	0847	GUANAJUATO	411	35	8.5%
64	0872	"	530	30	5.7%
65	0897	"	460	32	7.0%
66	1927	PENJAMO	447	30	6.7%
67	1928	"	426	38	8.9%
68	2004	"	201	37	18.4%
69	2125	SALAMANCA	578	20	3.5%
70	2167	"	610	30	4.9%
71	2192	"	602	30	5.0%
72	2193	"	847	60	7.1%
73	2228	"	549	30	5.5%
74	0357	CELAYA	251	30	12.0%
75	0366	"	1,190	40	3.4%
76	0410	"	646	31	4.8%

Tabla 1 (Continuación)

n	Número de la Sección	Municipio al que Pertenece	Votos Válidos en la Sección (M _i)	Votos Válidos Muestreados (m _i)	Proporción sobre total m _i / M _i
77	0423	CELAYA	427	30	7.0%
78	0462	"	233	35	15.0%
79	0464	"	342	30	8.8%
80	0466	"	702	30	4.3%
81	0510	"	327	32	9.8%
82	0515	"	73	30	41.1%
83	0555	"	292	30	10.3%
84	0966	IRAPUATO	493	31	6.3%
85	1015	"	434	28	6.5%
86	1018	"	497	30	6.0%
87	1032	"	529	31	5.9%
88	1070	"	480	27	5.6%
89	1075	"	434	30	6.9%
90	1104	"	375	30	8.0%
91	1105	"	717	30	4.2%
92	1145	"	281	32	11.4%
93	1678	LEON	450	40	8.9%
94	1803	"	881	32	3.6%
95	1703	"	229	41	17.9%
96	1778	"	409	30	7.3%
97	1335	"	556	30	5.4%
98	1404	"	479	30	6.3%
99	1853	"	607	30	4.9%
100	1578	"	380	30	7.9%
101	1529	"	654	30	4.6%
102	1729	"	582	20	3.4%
103	1554	"	972	30	3.1%
104	1852	"	651	29	4.5%
105	1627	"	599	25	4.2%
106	1652	"	356	36	10.1%
107	1741	"	860	19	2.2%
108	1330	"	741	30	4.0%
109	1380	"	269	30	11.2%
110	1429	"	356	30	8.4%
111	1454	"	278	30	10.8%
112	1479	"	297	30	10.1%
113	1728	"	453	30	6.6%
114	1505	"	438	30	6.8%
TOTALES			54,570	3,446	6.3%

4.1.3 Aplicación del Muestreo Bietápico

En Guanajuato la mayoría de las secciones se integraba cuando mucho por dos casillas: básica y contigua. Existieron algunas que estaban formadas por tres casillas lo cual significa que tanto la casilla básica como las contiguas poseían el mismo número. Las secciones están numeradas oficialmente del 1 al 3005 y las casillas suman un total de 4400.

Para la encuesta de salida las unidades de segunda etapa fueron aproximadamente treinta respuestas válidas obtenidas mediante el sistema de cuotas en cada sección; como los encuestadores no reportaron de cuántas casillas constaba la sección que se les asignó, en algunos cálculos se consideraran los votos válidos registrados por todas las casillas que conformaban la sección.

A continuación se enlistan las adecuaciones sobre las fórmulas de la sección 2.3.2 para aplicarlas de acuerdo a las particularidades de esta encuesta.

Las unidades primarias o de primera etapa serán los votos válidos reportados por cada sección (este dato se tomó del conteo rápido y, después, se verificó con los resultados oficiales que facilitó el IEE Guanajuato), y las secundarias o subunidades son las personas entrevistadas a la salida de la casilla que contestaron con validez por cual partido votaron.

M_i = Número de votos válidos en la i -ésima sección electoral, $i=1, 2, \dots, 3005$.

$\bar{M} = \frac{1}{N} \sum_{i=1}^N M_i$, promedio de votación válida por sección

\hat{p} = estimador de la proporción P de votos válidos pertenecientes a un partido político

N = Número de secciones totales (3005)

n = Número de secciones en muestra (114)

m_i = Número de votos válidos muestreados en la i -ésima sección muestral, $i=1, \dots, 114$

p_i = Proporción de votos válidos pertenecientes a algún partido en la i -ésima sección muestral

El estimador de la proporción para cada partido se obtendrá por medio de la fórmula 2.42 y el estimador de la varianza de esta misma proporción la dará la fórmula 2.44.

Los estimadores por intervalo para la proporción poblacional de cierta característica, para aquellas proporciones entre 0.3 y 0.7, se calculan mediante la siguiente expresión

$$\hat{p} \pm t \sqrt{\hat{V}(\hat{p})} \quad (4.1)$$

donde t es el desvío normal correspondiente al nivel de confianza.

4.1.4 Operativo de Campo

No sabemos cuántos equipos se designaron para el acopio de la información pero, operativamente se distribuyeron las cuadrillas en cada municipio donde se encontraban las secciones seleccionadas, quedando un encuestador por sección; en municipios donde sólo se observó una casilla, el levantador estaba equipado con teléfono celular; en municipios donde se observaron más casillas había un elemento que era el supervisor y hacía rondas en cada casilla para que le informaran los resultados y, éste a su vez, reportaba vía teléfono celular al CEO; por último, en municipios de las dimensiones de León hubo más de un supervisor y obviamente se asignaron más recursos humanos.

Una vez explicada la distribución de los recursos humanos, se describirá la operatividad del ejercicio: se realizaron tres intervalos de observación a cada sección, el primero de las 8:00 horas a las 10:30, el segundo de las 11:00 a las 14:30 horas y el tercero de 15:00 a 18:00 horas; se debían llenar cuotas de diez respuestas válidas en cada corte, teniendo entendido como respuesta válida la afirmación del votante hacia algún partido, si el votante declinaba la respuesta o afirmaba que había anulado su voto, se registraba de todos modos, pero no se consideraba como respuesta válida. De esta forma el operativo estimó tener capacidad suficiente para controlar con agilidad y calidad la información recogida en las 114 secciones muestreadas en el estado.

4.2 Resultados

4.2.1 Tamaño de Muestra

Tanto en la encuesta de salida como en el conteo rápido se pretendió tener un 4% aproximadamente del total de las secciones que cubren el estado; el porcentaje referido es de 120 secciones aunque, como ya se vio en la tabla 1, esto no fue posible y sólo se lograron registrar 114 secciones. El tamaño de muestra considerado obedece a los siguientes puntos.

- 1) Se consideró la información obtenida en la encuesta preelectoral que se analizó descriptivamente en el capítulo tres.
- 2) Dada la naturaleza de la información a registrar y la rapidez para concentrarla (conteo rápido y encuesta de salida), la logística y el operativo de campo son primordiales.

4.2.2 Preferencias

La tabla 2 muestra los resultados utilizando el diseño muestral, numéricamente se acercan más a los resultados oficiales que los de la encuesta preelectoral, una vez más, se hace énfasis en lo importante que es el momento en el tiempo en que se realiza una encuesta. Cada proporción arrojada por el muestreo bietápico se muestra con su respectiva varianza e intervalo de confianza. Por otro lado, la encuesta de salida no registra ningún cambio en el orden que ocuparon los partidos con respecto a la encuesta preelectoral.

ENCUESTA DE SALIDA			
TENDENCIAS CON MUESTREO BIFÁSICO			
PARTIDO	ESTIMADORES		INTERVALOS DE CONFIANZA
	PROPORCIÓN POBLACIONAL \hat{p}	VARIANZA DE LA PROPORCIÓN $\hat{var}(\hat{p})$	Al 95%
PAN	61.97	0.000532	(57.45, 66.49)
PRI	29.33	0.000443	(25.20, 33.46)
PRD	6.83	0.000123	(4.65, 9.01)
Otro	2.01	0.000026	(1.01, 3.01)

Tabla 2. Tendencias partidarias considerando muestreo bifásico aplicado sobre 3,446 respuestas válidas

En el resumen que se presenta en la tabla 3 se observan los resultados finales de la encuesta de salida tratados primero en forma descriptiva, después utilizando el muestreo bietápico, y al final se comparan con las proporciones oficiales.

GUANAJUATO			
COMPARATIVO GLOBAL TENDENCIAS - OFICIALES			
ENCUESTA DE SALIDA (Mayo 1995)			
PARTIDO	SIN DISEÑO	CON DISEÑO	OFICIAL
PAN	61.09	61.97	58.10
PRI	29.19	29.33	32.90
PRD	7.75	6.83	7.02
Otro	1.97	2.01	1.98

Tabla 3. Resultados globales de la encuesta de salida, (sin y considerando diseño de muestreo sobre una población de 3,446 votos válidos) y resultados oficiales (IEE Guanajuato, Elección extraordinaria para gobernador 1995. Sobre 1,244,955 votos válidos)

4.2.3 Preferencias por Partido en cada estrato

El propósito de esta sección es únicamente descriptivo e ilustrativo de las tendencias recabadas por la encuesta de salida en los estratos que se utilizaron en el capítulo anterior.

Aunque los estratos son los mismos, la muestra de municipios extraída de cada estrato no es igual a la muestra utilizada en la encuesta preelectoral. La figura 1 muestra los resultados.

Dado que la tabla 26 del capítulo tres mostró que el PRI ganó en el Estrato 1, se antoja interesante observar si la encuesta de salida registra lo mismo; pero la figura 1 no refleja nada al respecto, de hecho, el estrato donde alcanza su valor máximo el PRI es en el Estrato 2.

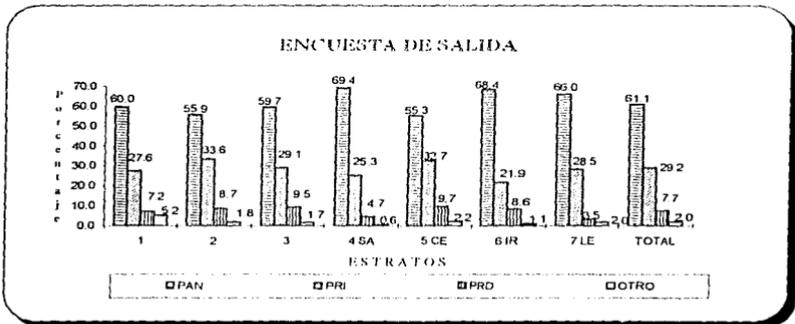


Figura 1. Tendencias en cada estrato registradas por la encuesta de salida (sin diseño muestral). 3446 respuestas

CAPITULO 5 Conteo Rápido

5.1 Introducción

El conteo rápido es el tema de este capítulo. En primer término se hace una descripción del procedimiento, el método de muestreo empleado y el operativo de campo. En el segundo apartado, llamado resultados se da el tamaño de muestra, el desenlace derivado del cierre de las casillas muestreadas, y finalmente, las preferencias en los estratos conocidos.

5.1.1 Descripción

El mecanismo del conteo rápido es sencillo en cuanto a la operatividad, porque el informador debe esperar a que se cierre la casilla para registrar los resultados, y comunicarlos o enviarlos a un centro de acopio de información.

El objetivo fundamental de este procedimiento es comunicar la información final con la mayor rapidez posible para que, también lo más pronto posible, se procese; y así obtener tendencias estadísticamente confiables sobre las elecciones, a pesar de que aún no se cuente con la totalidad de los resultados.

Un factor importante en el desarrollo de este ejercicio, qué en la práctica debería realizarse, es el ensayo, porque es de gran utilidad para minimizar los imprevistos que siempre existen durante la ejecución del proceso; este factor debe de estar a punto por lo menos un día antes de la celebración de los comicios. La gran desventaja del ensayo es su costo elevado.

El aspecto de la comunicación también es muy importante, porque se pueden encontrar regiones geográficas que por su orografía impidan la salida de comunicación vía aire (telefonía celular, ondas de radio, microondas etc.) o, simplemente que en la región muestreada no existan este tipo de medios, por lo que, el enumerador tiene que improvisar para comunicar el cierre de sección a la brevedad posible, pero la mayoría de las veces esto resulta bastante costoso en lo que a tiempo se refiere, que es lo más valioso en el conteo rápido.

Otro elemento importante a considerar es el humano; los informadores deben ser responsables y honestos, conscientes de que tienen que cumplir con su deber y no hacer mal uso del cierre de sección, porque la calidad de información que se maneja en los conteos rápidos es muy delicada; y es preferible perderla, a codificarla si es que ésta es equivocada. Esto exige que debe existir un proceso de selección de los enumeradores muy riguroso, y así reducir las posibilidades de procesar información errónea.

Por último, no podemos pasar por alto que debe existir un instrumento de contingencia para sustituir las casillas no localizadas o que hayan reportado problemas en cuanto a la transmisión del cierre.

5.1.2 Diseño Muestral

El muestreo por conglomerados se describió anteriormente pero, como también se utilizó muestreo sistemático con arranque aleatorio, es necesaria una breve explicación sobre la relación que existe entre ambos.

En forma general y suponiendo que N se puede expresar por el producto de k y n , se debe extraer un número al azar que sea menor o igual a k , denotado por i , de acuerdo a la forma de elegir las siguientes unidades explicada en 4.1.2; se seleccionará cada k -ésima unidad de la población por lo tanto, la muestra contendrá n unidades etiquetadas con números ordenados i , es decir, $i, i+k, i+2k, \dots, i+(n-1)k$.

La relación con el muestreo por conglomerados consiste en que al poder expresar a N como un producto nk es posible descomponer a la población en k unidades de muestra, cada una constituida por n unidades, de tal forma que se puedan acomodar en un arreglo matricial donde cada columna representa una de las k muestras y los renglones son los elementos n que le dan forma; por lo tanto, si sólo se elige al azar una de esas k muestras se obtiene una sola unidad de muestreo que en sí misma es una muestra total. Entonces, de acuerdo a la descripción del muestreo sistemático explicada en el capítulo anterior, se debe asumir que las k muestras son k unidades conglomeradas, y que es posible aplicar un muestreo aleatorio simple sobre éstas, lo que resulta ser una muestra sistemática.

De esta forma, cada sección representa a un conglomerado de electores. Por lo tanto, el total del electorado (lista nominal) se encuentra agrupado o conglomerado en secciones electorales.

En la tabla 1 se presenta la distribución de la muestra para el conteo rápido. Existieron varios problemas que impidieron el censo de algunas secciones, la solución inmediata fue sustituirlas con todas las reservas necesarias en cada caso, pero de esto se hablará en la sección 5.1.4.

La tabla 1 describe de izquierda a derecha; el número ascendente de secciones seleccionadas en la muestra (tanto originales como sustitutas), el número de sección asignado por el IFE, el municipio al que pertenece, los votos válidos y la lista nominal oficial y, finalmente, la participación (votos válidos entre lista nominal).

5.1.3 Aplicación del Muestreo por Conglomerados

Como ya se explicó en la sección 4.1.3 las secciones en Guanajuato estaban formadas por un máximo de tres casillas, así que la unidad mínima de muestreo en el conteo rápido fue la sección.

Recordando que en el muestreo por conglomerados es posible escoger entre tomar una muestra del conglomerado o censarlo, en el caso del conteo rápido se está censando porque se registran los resultados definitivos de cada sección.

CONTEO RÁPIDO					
DISTRIBUCIÓN DE LA MUESTRA					
n	Número de Sección	Municipio al que Pertenece	Votos Válidos en las Secciones	Lista Nominal de las Secciones	Participación en las Secciones
1	2736	TARANDACUAO	309	725	42.62
2	0750	DR. MORA	310	552	56.16
3	0922	HUANIMARO	603	1,076	56.04
4	2026*	PUEBLO NUEVO	623	950	65.58
5	2935*	XICHU	352	732	48.09
6	1894*	OCAMPO	431	683	63.10
7	2891*	VICTORIA	622	894	69.57
8	2892*	"	626	921	67.97
9	0725	CUERAMARO	342	631	54.20
10	2350	S.D. DE LA UNION	431	698	61.75
11	2054	PURISIMA DEL R.	320	511	62.62
12	1183	JARAL DEL PROG.	612	1,255	48.76
13	0575	CD. M. DOBLADO	544	814	66.83
14	2920	VILLAGRAN	276	496	55.65
15	2100	ROMITA	161	340	47.35
16	0250	APASEO EL ALTO	364	713	51.05
17	0254*	"	411	752	54.65
18	2796*	URIANGATO	568	863	65.82
19	1233	JEREQUARO	184	444	41.44
20	1208	"	368	727	50.62
21	1204*	"	331	625	52.96
22	1878	MOROLEON	463	745	62.15
23	0625	COMONFORT	254	436	58.26
24	0617*	"	672	1,165	57.68
25	2625*	STA. CRUZ DE J.R.	759	1,170	64.87
26	0296*	APASEO EL GDE.	658	971	67.77
27	0297*	"	545	825	66.06
28	0050	ABASOLO	470	788	59.64
29	0017*	"	461	787	58.58
30	2545*	S. LUIS DE LA PAZ	647	1,073	60.30
31	2553*	"	746	1,204	61.96
32	0675	CORTAZAR	449	693	64.79
33	0663*	"	502	783	64.11
34	2377*	SAN FELIPE	424	714	59.38
35	2386*	"	436	794	54.91
36	2450	S. FCO. DEL RIN.	488	707	69.02
37	2475	"	349	551	63.34
38	2500	"	212	434	48.85
39	0762	DOLORES HGO.	765	1,296	59.03

Tabla 1. Distribución muestral para el Cuento Rápido.

Tabla 1 (Continuación)

n	Número Sección	Municipio al que Pertenece	V. Válidos (M.)	Lista Nominal	Participación %
40	0757*	DOLORES HGO.	530	852	62.21
41	0763*	"	569	882	64.51
42	2275	SALVATIERRA	446	955	46.70
43	2270*	"	528	820	64.39
44	2269*	"	458	894	51.23
45	0150	ALLENDE	508	730	69.59
46	0153*	"	709	1,020	69.51
47	0151*	"	439	630	69.68
48	0162*	"	370	532	69.55
49	2668	SILAO	665	987	67.38
50	2663	"	772	1,237	62.41
51	0075	ACAMBARO	322	657	49.01
52	0056*	"	566	1,124	50.36
53	0062*	"	478	865	55.26
54	0897	GUANAJUATO	460	691	66.57
55	0872	"	530	925	57.30
56	0847	"	411	704	58.38
57	2820	VALLE DE SNTGO.	501	727	68.91
58	2845	"	377	784	48.09
59	2838*	"	510	985	51.78
60	2823*	"	701	1,184	59.21
61	1928	PENJAMO	426	698	61.03
62	2004	"	201	435	46.21
63	1927*	"	447	836	53.47
64	1926*	"	681	1,013	67.23
65	2192	SALAMANCA	602	909	66.23
66	2125	"	578	818	70.66
67	2167	"	610	971	62.82
68	2228	"	549	767	71.58
69	2193*	"	847	1,480	57.23
70	0555	CELAYA	292	551	52.99
71	0510	"	327	661	49.47
72	0410	"	646	1,015	63.65
73	0462	"	233	373	62.47
74	0366	"	1,190	1,874	63.50
75	0464	"	342	535	63.93
76	0423	"	427	724	58.98
77	0515	"	73	102	71.57
78	0357	"	251	444	56.53

Tabla 1 (Continuación)

n	Número Sección	Municipio al que Pertenece	V. Válidos (M.)	Lista Nominal	Participación %
79	0966	IRAPUATO	493	874	56.41
80	1145	"	281	513	54.78
81	1018	"	497	754	65.92
82	1075	"	434	760	57.11
83	1070	"	480	885	54.24
84	1104	"	375	634	59.15
85	1015	"	434	700	62.00
86	1105	"	717	1,212	59.16
87	1032*	"	529	794	66.62
88	1505	LEON	438	690	63.48
89	1728	"	453	576	78.65
90	1380	"	269	373	72.12
91	1652	"	356	494	72.06
92	1429	"	356	520	68.46
93	1404	"	479	700	68.43
94	1454	"	278	393	70.74
95	1529	"	654	863	75.78
96	1678	"	450	734	61.31
97	1330	"	741	1,165	63.61
98	1853	"	607	889	68.28
99	1803	"	881	1,380	63.84
100	1578	"	380	623	61.00
101	1778	"	409	600	68.17
102	1703	"	229	313	73.16
103	1335	"	556	868	64.06
104	1554	"	972	1,534	63.36
105	1729	"	582	851	68.39
106	1627	"	599	1,257	47.65
107	1741	"	860	1,278	67.29
T O T A L E S			52,449	86,156	60.88
* Secciones sustitutas 74 secciones originales 33 secciones sustitutas					

La notación para el censo rápido se obtiene mediante la combinación y adaptación de las ecuaciones contenidas en las secciones 1.2.5, 2.2.2 y 2.3.2 cuyas reformulaciones y nuevas expresiones se presentan en seguida.

\hat{p} = estimador de la proporción P de personas que votaron por un partido específico

N = número de secciones o conglomerados totales (3005)

n = número de secciones o conglomerados en muestra (74 sin sustituir y 107 con sustitución)

M_i = número de votantes de la sección electoral i -ésima, $i = 1, \dots, N$ (votos válidos)

m_i = número de votantes de la sección electoral muestral i -ésima, $i = 1, \dots, n$

Como las unidades de muestreo son conglomerados, el estimador que se utiliza es la siguiente adecuación de la fórmula (2.42) del capítulo 2

$$\hat{p} = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n M_i} \quad (5.1)$$

donde y_i es el número de personas que votó por un partido específico en la sección i , $i = 1, \dots, n$.

La varianza se obtiene de la fórmula (2.43) y se expresa como sigue

$$V(\hat{p}) = \frac{N-n}{Nn} S_b^2 \quad (5.2)$$

donde S_b^2 es la ecuación siguiente

$$S_b^2 = \frac{1}{N-1} \sum_{i=1}^N \frac{M_i^2}{\bar{M}^2} (p_i - P)^2 \quad (5.3)$$

en donde \bar{M} es

$$\bar{M} = \frac{1}{N} \sum_{i=1}^N M_i \quad (5.4)$$

y p_i es la proporción de votantes que se inclinaron por un partido en la sección i .

Un estimador de S_b^2 es \hat{S}_b^2 cuya expresión es

$$\hat{S}_b^2 = \frac{1}{n-1} \sum_{i=1}^n \frac{M_i^2}{\bar{M}^2} (p_i - \hat{p})^2 \quad (5.5)$$

donde utilizando (2.46)

$$\bar{M}_n = \frac{1}{n} \sum_{i=1}^n M_i \quad (5.6)$$

por lo tanto, un estimador de (5.2) se obtiene si se sustituye \hat{S}_b^2 por S_b^2 , con esta modificación sustentada en (2.44) y (2.45), se obtiene la ecuación del estimador de la varianza

$$\hat{V}(\hat{p}) = \frac{N-n}{Nn} \hat{S}_b^2 \quad (5.7)$$

Los estimadores por intervalo para la proporción poblacional de cierta característica, para aquellas proporciones entre 0.3 y 0.7, se calculan mediante la siguiente expresión

$$\hat{p} \pm t \sqrt{\hat{p}(1-\hat{p})} \quad (5.8)$$

donde t es el desvío normal correspondiente al grado de confianza deseada.

Referencia: Sukhatme (1984, pp. 323-336)

5.1.4 Operativo de campo

El operativo del cuento rápido fue el que sufrió mayores contratiempos; hasta cierto punto estas complicaciones son normales. Como el enemigo principal del cuento rápido es el desperdicio de tiempo. La paradoja en los conteos rápidos, es que si surge un problema, se necesita tiempo para tomar la decisión que lo resolverá y precisamente es de lo que más se carece.

El problema principal en los comicios de Guanajuato fue no haber registrado algunas de las secciones previamente elegidas por el muestreo sistemático, y tampoco se contaba con un plan diseñado especialmente (desde un punto de vista muestral) para realizar sustituciones. Esto provocó que se hicieran sustituciones; en muchos casos el criterio común para sustituir fue trasladarse a la sección más cercana a la asignada. Después de todo, en el aspecto del diseño y para los efectos de este trabajo, se decidió separar los resultados en dos clases: secciones originales y su complemento (llamadas secciones sustitutas).

Para la realización del cuento rápido se utilizó la misma selección de la encuesta de salida. Sin embargo, si se comparan las tablas de distribución muestral de la encuesta de salida y del cuento rápido, se notará que el municipio Yuriria no aparece en la distribución muestral del cuento rápido (tabla 1 de este capítulo).

La mayoría de los registros de cierre de casilla fueron correctos pero, se encontraron dos tipos de error, que se pudieron corregir al compararlos con los resultados oficiales desagregados del IEE de Guanajuato. El primer error y el más común fue que en las secciones que tenían dos casillas (básica y contigua) sólo se tomaba el cierre de una, esto denotaba un nivel de participación muy alto. El segundo error, que afortunadamente solamente se detectó en una sección, fue información distinta a la del cierre de la casilla. Existieron otros errores no atribuibles al diseño muestral que fueron los que provocaron la mayoría de las sustituciones, estos se expondrán en el apartado A-1 del anexo.

5.2 Resultados

5.2.1 Tamaño de Muestra

En la sección 4.2.1 se explicó que el tamaño de muestra calculado era el 4% aproximadamente del total de las secciones electorales, la muestra logró un total de 107 secciones censadas considerando casillas sustitutas, esto es el 3.56% de total de 3005; además,

las listas nominales acumuladas de estas secciones representan el 3.98% de la lista nominal de todo el estado. Sin considerar las casillas sustitutas se reportaron 74 secciones, que son el 2.46% de las 3005, cuyo acumulado nominal es el 2.60% del total estatal. Los puntos en los que se basó el tamaño de la muestra son los mismos que están expuestos en la sección 4.2.1 del capítulo anterior. Para la realización del muestreo sistemático se tomaron las secciones cada 25 posiciones según estaban acomodadas en el encarte.

5.2.2 Tendencias

Conviene recordar que los resultados se presentarán en el siguiente orden: primero sin sustituciones y luego con sustituciones; independientemente de esta clasificación se presentará un comparativo de los dos tipos de resultados para verificar si existieron grandes diferencias.

Aprovechando que se tenían los datos muy a la mano, con el propósito de ilustrar lo parecido que fueron los valores numéricos entre secciones originales y sustitutas; se realizaron algunos cálculos en forma descriptiva que no aparecen en las tablas de este capítulo. El promedio de participación de las secciones originales es 60.87% y si adicionamos las secciones sustitutas, el promedio resultante es 60.94%. La participación global que reflejan las secciones originales es del 60.83%, mientras que utilizando todas se obtuvo un 60.88%, (son casi iguales). La participación oficial fue de 57.55%.

Las tablas 2 y 3 exhiben (secciones originales y sustitutas respectivamente) los resultados obtenidos con muestreo por conglomerados son: los estimadores proporcionales, la varianza de cada uno y los intervalos de confianza correspondientes. Los cálculos de la varianza y los intervalos de confianza se realizaron con las expresiones asociadas de la sección 5.1.3.

CONTEO RÁPIDO (SIN SUSTITUCIÓN)			
TENDENCIAS CON MUESTREO POR CONGLOMERADOS			
PARTIDO	ESTIMADORES		INTERVALOS DE CONFIANZA
	ESTIMADOR DE PROPORCIÓN	VARIANZA	Al 95%
	\hat{p}	$\hat{v}ar(\hat{p})$	
PAN	62.1	0.000261	(58.93, 65.27)
PR1	29.0	0.000165	(26.48, 31.51)
PRD	7.0	0.000083	(5.22, 8.78)
Otro	1.9	0.000004	(1.49, 2.31)

Tabla 2. Tendencias partidarias con diseño muestral (por Conglomerados descrito la sección 5.1.2) sin secciones sustitutas

CONTEO RÁPIDO (CON SUSTITUCIÓN)			
TENDENCIAS CON MUESTREO POR CONGLOMERADOS			
PARTIDO	ESTIMADORES		INTERVALOS DE CONFIANZA
	ESTIMADOR DE PROPORCIÓN	VARIANZA	Al 95%
	\hat{p}	$\hat{var}(\hat{p})$	
PAN	60.84	0.000154	(58.41, 63.27)
PRI	30.04	0.000119	(27.91, 32.18)
PRD	7.38	0.000060	(5.86, 8.90)
Otro	1.74	0.000002	(1.44, 2.04)

Tabla 3. Tendencias partidarias con diseño muestral (por Conglomerados descrito la sección 5.1.2) con secciones sustitutas

Ahora se presenta la tabla 4 que contiene los resultados de ambas clasificaciones y los resultados oficiales. Es importante observar que numéricamente son muy parecidas ambas clasificaciones.

GUANAJUATO			
COMPARATIVO GLOBAL TENDENCIAS - OFICIALES			
CONTEO RÁPIDO (Mayo 1995)			
PARTIDO	Sin Sustitución	Con Sustitución	Oficial
PAN	62.10	60.84	58.10
PRI	29.00	30.04	32.90
PRD	7.00	7.38	7.02
Otro	1.90	1.74	1.98
Participación	60.83	60.88	57.55
Votos Válidos	34,273	52,449	1,244,955
Secciones	74	107	3005

Tabla 4. Resultados globales del conteo rápido, (sin y considerando sustituciones) y resultados oficiales (IEE Guanajuato, Elección extraordinaria para gobernador 1995)

5.2.3 Tendencias por cada Estrato

Una vez más, y exclusivamente con fines descriptivos, se presentan las tendencias que arrojó el conteo rápido en los estratos utilizados en el capítulo tres, la figura 1 no contempla las secciones sustitutas y la figura 2 sí.

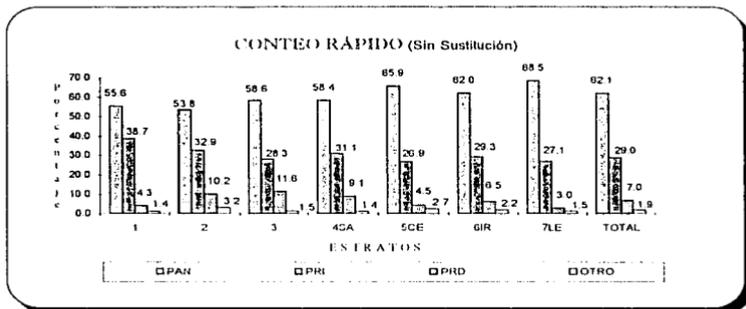


Figura 1. Tendencias en cada estrato registradas por el conteo rápido (sin diseño muestral). Se censaron 74 secciones

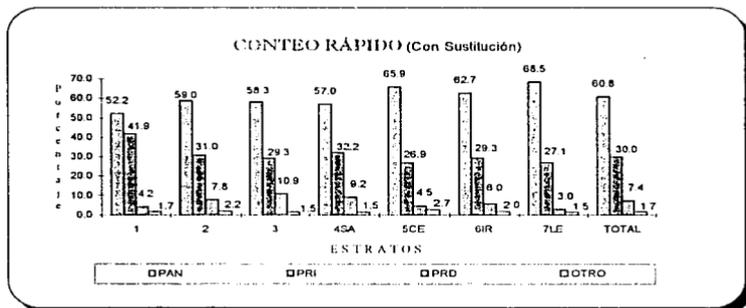


Figura 2. Tendencias en cada estrato registradas por el conteo rápido (sin diseño muestral). Se censaron 107 secciones

Recordando el seguimiento especial sobre el estrato uno, se observa que tanto la figura 1 como la 2 muestran que el PRI alcanza su máximo valor en este estrato; Aún así, estaban lejos de las proporciones alcanzadas por el PAN. Por otro lado, las elecciones eran a nivel estatal y el estrato en particular, representa el 3.61% de la población potencial con derecho a voto de todo el estado, de acuerdo al código 90 del INEGI.

Para finalizar, aunque los estimadores de las proporciones alcanzadas por el PAN sobrepasaron ligeramente la proporción real, sí dieron un claro indicio sobre el partido ganador de la contienda electoral.

CAPITULO 6 Conclusiones

Desde el sexenio 1988-1994 se han revelado en forma más abierta los resultados de encuestas de opinión dirigidas a cuestiones electorales.

Ahora con las reformas recientes al Código Federal de Instituciones y Procedimientos Electorales (COFIPE) este tipo de estudios de opinión sufren una transición de tabú a asunto de dominio público, incluso se convierten en requisito necesario (no por ley, pero sí por costumbre) para robustecer la credibilidad de la transparencia de un proceso electoral.

Es así cuando en 1988 en las elecciones federales para presidente de la República Mexicana aparece el primer intento de aplicar un mecanismo postelectoral cuyo objetivo era tener mediante una muestra, tendencias muy confiables para dar a conocer al candidato triunfador. Como se recuerda, el sistema de cómputo utilizado se cayó y no fue posible aplicar el procedimiento en cuestión.

Tal vez el primer conteo rápido oficial fue precisamente el utilizado en las elecciones federales para presidente de los Estados Unidos Mexicanos correspondientes al sexenio 1994-200. Después de éste se practicaron más en las diversas elecciones ordinarias y extraordinarias de carácter municipal y estatal.

Este fenómeno de apertura electoral tiene como consecuencia que se abra en México un nuevo nicho en el campo de la estadística aplicada al muestreo. Porque de sólo hacer estudios de opinión dirigidos a la rama de la mercadotecnia (y aunque se hayan realizado algunas de carácter electoral, de todos modos eran información privada), ahora se pueden aplicar al ramo electoral y desarrollar diseños muestrales específicos para este tipo de encuestas y, lo más importante, poder divulgarlas con amplia libertad.

Actualmente en las empresas de México, que desde tiempo atrás se dedicaban a realizar encuestas de opinión, han surgido nuevas divisiones que se dedican exclusivamente a los actos electorales. También se han constituido en el país nuevas empresas y organismos especializados en encuestas electorales e incluso en consultoría sobre este tema.

Como se mencionó en la introducción de este trabajo los diseños muestrales combinan diferentes técnicas de muestreo. Sobre estas técnicas ya no hay mucho que descubrir, aún así, continúa la investigación a nivel teórico. Es así que el verdadero reto para el muestrista en la práctica es lograr construir un diseño muestral que cumpla con los objetivos requeridos por una encuesta electoral.

Para la elaboración de la encuesta preelectoral correspondiente a las elecciones extraordinarias a gobernador de Guanajuato del 28 de mayo de 1995, se combinaron: el muestreo aleatorio simple, el muestreo estratificado, el muestreo bietápico y el muestreo por cuotas. Un objetivo general de la encuesta preelectoral fue el determinar el nivel de información y comportamiento de la población con derecho a voto con respecto a la jornada electoral. Considero que en este sentido se obtuvo información muy limpia, porque las proporciones de no respuesta fueron bajas en casi todas las preguntas, lo que demuestra un aceptable grado de

información; la alta intención de participar del electorado de alguna forma señala el comportamiento del mismo.

Uno de los objetivos específicos de la encuesta preelectoral era generar información para saber el nivel en un escalafón y las tendencias de cada opción política en ese momento, otro el de estimar el nivel de algunas variables sociodemográficas. En cuanto al primer objetivo pienso que se logró; aunque, es fundamental el punto en el tiempo donde se realizó la encuesta. No se debe olvidar que la toma de decisiones en una elección es muy dinámica y que de un instante a otro varía; quizá el punto débil en este objetivo, consistió en que no fue muy atinado brindar dos opciones que no indicaran una respuesta clara sobre alguno de los partidos. Por ejemplo, al preguntar por qué partido votaría, se le dan todas las opciones políticas y la de ninguno, pero no se le da la oportunidad de escoger: No he decidido; si él contesta de inmediato que no lo sabe o que aún no decide entonces, así se registra.

En la encuesta, la opción **No he decidido** superó a la considerada tercera fuerza política en el estado, es natural que suceda así en una encuesta preelectoral, pero también es deseable tener la seguridad de que si un elector contesta que no sabe por quién votar, esta respuesta sea verdadera.

En relación al objetivo de las variables sociodemográficas, tal vez no se cumplió como se debía, porque las proporciones no fueron cubiertas como se esperaba debido a que las cuotas asignadas no se llenaron cabalmente. Es aquí donde se comprende la polémica que el muestreo por cuotas ha generado entre la comunidad estadística, porque, por un lado lo califican de pseudoaleatorio y por otro de lo contrario, además de ser muy práctico y barato. El calificativo de pseudoaleatorio es porque depende totalmente de la subjetividad del encuestador y del riesgo latente de un sesgo producido por el incumplimiento de los controles y esquemas de las cuotas (ver tesina Tres Encuestas por Muestreo, 1996 pp. 25-26).

El muestreo por cuotas provocó en la encuesta preelectoral cuotas sobresaturadas, lo que dio como resultado en la variable sexo, mayor proporción masculina, cuando lo normal hubiera sido (de acuerdo a datos del INEGI) mayor cantidad de mujeres. En la variable edad también existieron algunas irregularidades. No va a ser posible identificar qué causó el problema de las variables cuota; pero, repito que se debe tener más cuidado en el entrenamiento de los encuestadores, para que comprendan los controles y se apeguen lo mejor posible a lo que se les pide. Retomando las inconsistencias que se presentaron en la variable edad, quizá el problema fue la manera de estructurar los grupos de edad y no la de los controles de cuota. Y hablando de la estructuración de las categorías de las variables, si se va a aprovechar una encuesta preelectoral para profundizar en las características sociodemográficas del electorado, lo más prudente es homologar las categorías de cada variable a las ya establecidas por organismos reconocidos en el país.

Esto lo manifiesto porque hubiera sido muy interesante poder comparar los resultados del INEGI con los recabados por la encuesta de una forma directa. Incluso también se pudo haber sentado el precedente sociodemográfico particular sobre la población con derecho a voto en el estado. Pero fueron tan incompatibles algunas categorías de las variables que sería imposible

establecer relaciones estadísticamente respaldadas entre variables sociodemográficas y variables electorales.

La encuesta de salida en términos generales funcionó bien, la mayoría de las cuotas de electores fueron satisfechas. En encuestas de salida no se atribuye exclusivamente la causa de las fallas a los encuestadores; existen otras áreas donde pueden ocurrir omisiones, en el caso del conteo rápido y la encuesta de salida, las omisiones de información pueden ocurrir en el momento de la transmisión o la captura. En la encuesta de salida se encontraron defectos como los siguientes ejemplos: hubo secciones en las que no se practicaron todos los cortes requeridos esto se notaba porque el número de sección aparecía sólo una vez. Lo correcto es que dicho número debía aparecer tres veces, esto indicaba que se habían realizado los tres cortes. También sucedió lo contrario, es decir, un número de sección llegó a repetirse hasta en cuatro ocasiones.

La encuesta de salida se vio afectada sobretudo en el aspecto de la selección sistemática, porque hubo un 31% de sustituciones, la mayoría provocadas por causas ajenas al diseño. La causa más frecuente que provocó estas sustituciones fue que no dejaron cumplir con su labor a los encuestadores, incluso ocasionalmente se reportaron arrestos y encarcelamientos; en el anexo se mencionan en forma específica otras causantes de las sustituciones.

El conteo rápido corrió ciertos riesgos, pero todo indica que se superaron y su funcionamiento fue correcto; la mayoría de los problemas fueron provocados por cuestiones no relacionadas con el muestreo. El conteo rápido utilizó muestreo por conglomerados y selección sistemática. Como las secciones en muestra son las mismas tanto para conteo rápido como para encuesta de salida, el conteo rápido soportó las sustituciones, aunque las reglas para sustituir fueron menos flexibles que las de la encuesta de salida (esto también se explica en el anexo).

Por estratos, el número 1 no tuvo el mismo comportamiento que los demás. Al tomar este estrato completo dentro del marco de los resultados oficiales observamos que el partido ganador es el PRI con un pequeño margen de ventaja sobre el PAN.

El estrato 1 en las encuestas y el conteo siempre estuvo dominado por el PAN, aunque en la preelectoral y en el conteo rápido si se refleja un estrechamiento en la ventaja del PAN sobre el PRI de 10 puntos porcentuales aproximadamente (tabla 27 cap. 3 y figura 2 cap. 5).

El estrato 1 aglomeró a aquellas poblaciones que tenían menos de 10,000 habitantes con derecho a voto, que de acuerdo al INEGI contienen en su mayoría localidades rurales. El IEE Guanajuato lo confirma al clasificar la mayoría de las casillas ubicadas en estos municipios como rurales. Este comportamiento es parecido a lo que sucedió en las elecciones a gobernador de Michoacán en noviembre de 1995, en donde las encuestas practicadas por el CEO mostraron que el PRI fue derrotado en los municipios con mayor urbanización, o que contenían las principales ciudades; tales como la capital (J. L. Castrejón, Tres Encuestas por Muestreo, 1996, pp. 39 y 53). Lo que hizo triunfar al PRI fue su simpatía en municipios menos urbanizados y en comunidades rurales. Tratando de encontrar una explicación a este suceso en Guanajuato, es posible considerar que el tipo de población que aglomera el estrato en cuestión posee características similares a las localidades en las que el PRI triunfó en Michoacán. Por otra parte el estrato 1 participa con el 3.61% de la población potencial con derecho a voto.

Por otra parte, es recomendable la elaboración de un programa formal de sustitución de casillas con reglas básicas para que su aplicación se lleve a cabo sistemática y velozmente respetando los lineamientos de un conteo rápido.

Con respecto a los resultados de las elecciones: Guanajuato era uno de los estados con gobernador interino, por lo que se convocó a elecciones extraordinarias para gobernador el 28 de mayo de 1995. Todo lo que se puede decir es que el resultado es claramente unipartidista.

A continuación, se desglosan en las figuras 1, 2, 3 y 4 los resultados totales por partido según los estratos en cada uno de los ejercicios practicados por el CEO y los resultados oficiales; la figura 5 muestra el resumen de las cuatro anteriores.

Las características de las figuras están basadas en la siguiente información:

ENCUESTA PREELECTORAL ➤ Calculada sobre 1,298 respuestas válidas.

ENCUESTA DE SALIDA ➤ Calculada sobre 3,446 respuestas válidas de 114 secciones muestreadas.

CONTEO RÁPIDO ➤ Calculado sobre 52,449 votos válidos producto de 107 secciones censadas. El total de secciones es de 3,005.

RESULTADOS OFICIALES ➤ Calculados sobre 1,244,955 votos válidos y tomando la totalidad de los municipios que formaban cada estrato. La lista nominal del estado es de 2,163,111 ciudadanos.

ESTRATOS	CARACTERÍSTICAS
1	Aglomeró municipios cuya población mayor a 18 años de acuerdo a código 90 fueran menores a 10,001 habitantes. Este estrato representa el 3.61% de la población mayor de 18 años de todo el estado. ¹
2	Aglomeró municipios cuya población mayor a 18 años de acuerdo a código 90 fueran desde 10,001 hasta 30,000 habitantes. Este estrato representa el 16.72% de la población mayor a 18 años de todo el estado.
3	Aglomeró municipios cuya población mayor a 18 años de acuerdo a código 90 fueran desde 30,001 hasta 70,000 votantes. Este estrato representa el 34.22% de la población mayor a 18 años de todo el estado.

4 (SALAMANCA)	Población mayor de 18 años según INEGI: 109,438 habitantes. Representa el 5.49% de la población mayor a 18 años de todo el estado.
5 (CELAYA)	Población mayor de 18 años según INEGI: 163,012 habitantes. Representa el 8.17% de la población mayor a 18 años de todo el estado.
6 (IRAPUATO)	Población mayor de 18 años según INEGI: 187,427 habitantes. Representa el 9.40% de la población mayor a 18 años de todo el estado.
7 (LEÓN)	Población mayor de 18 años según INEGI: 446,299 habitantes. Representa el 22.38% de la población mayor a 18 años de todo el estado.
<p>La población mayor de 18 años del estado de Guanajuato según código 90 del INEGI es de 1,994,161 habitantes.</p>	

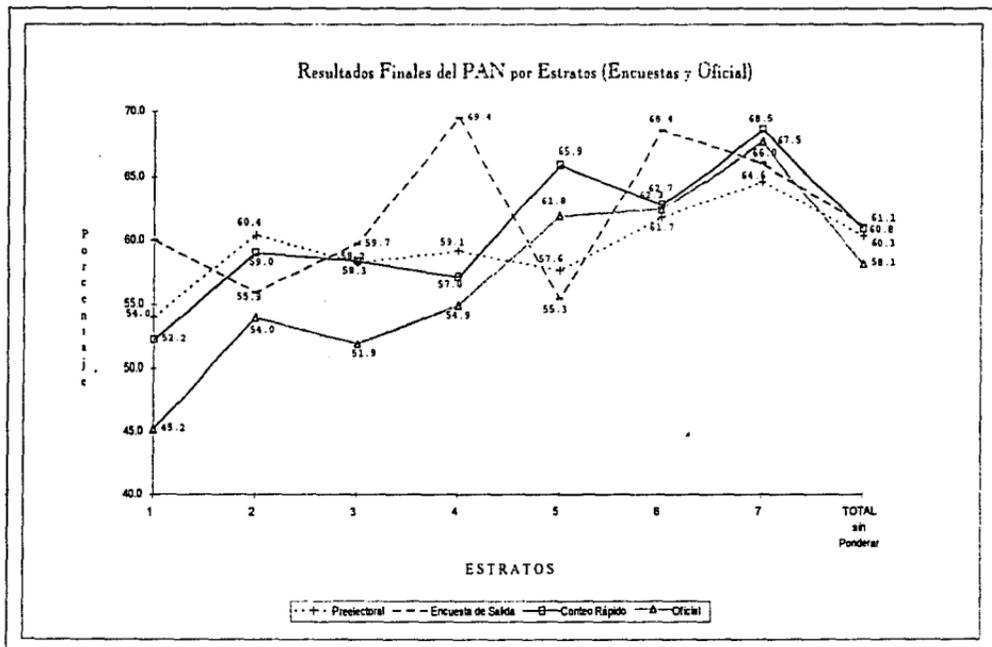


Figura 1. Resultados obtenidos por el Partido Acción Nacional en las encuestas preelectoral y de salida, en el censo rápido y los resultados oficiales por estratos

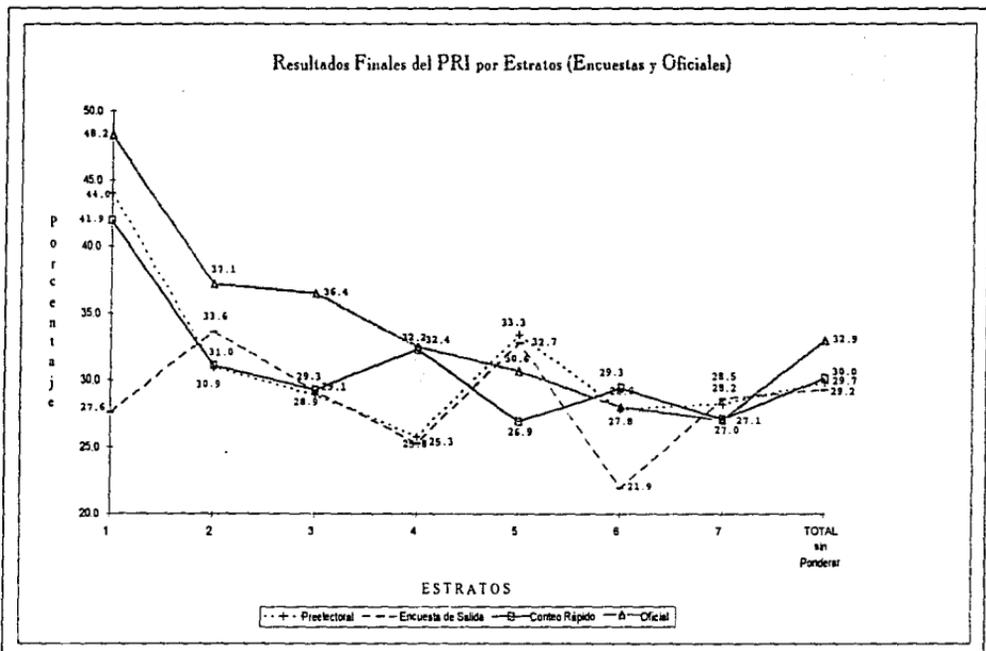


Figura 2. Resultados obtenidos por el Partido Revolucionario Institucional en las encuestas preelectoral y de salida, en el censo rápido y los resultados oficiales por estratos

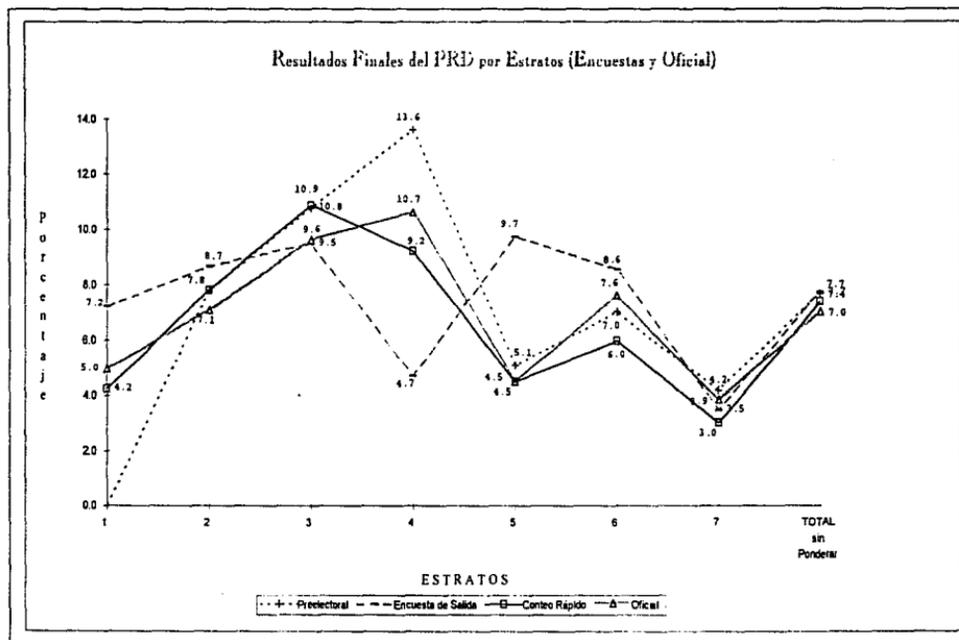


Figura 3. Resultados obtenidos por el Partido de la Revolución Democrática en las encuestas preelectoral y de salida, en el censo rápido y los resultados oficiales por estratos

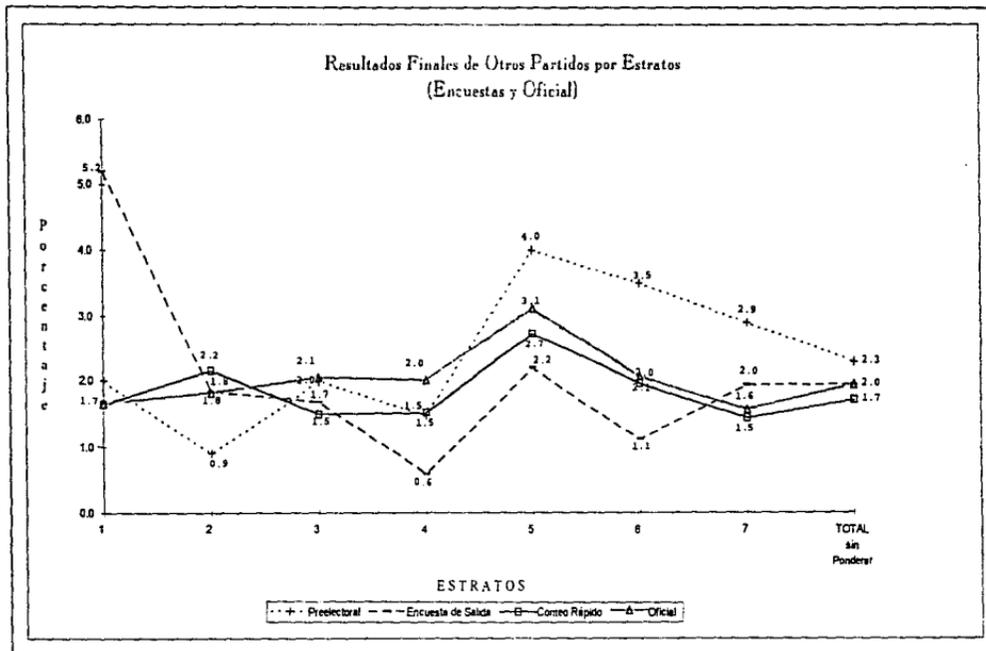


Figura 4. Resultados obtenidos por otros partidos en las encuestas preelectoral y de salida, en el conteo rápido y los resultados oficiales por estratos

Preelectoral, Encuesta de Salida, Conteo Rápido y Resultados Oficiales

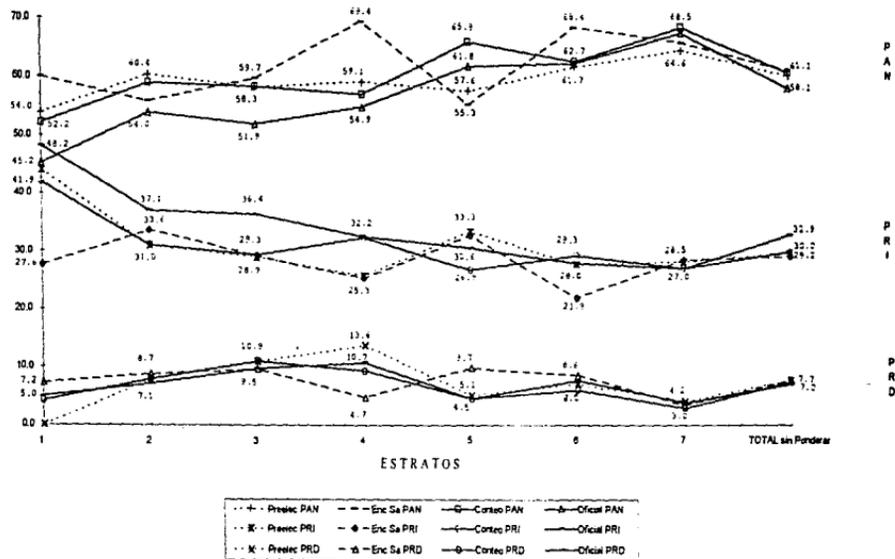


Figura 5. Resumen comparativo de los tres mecanismos aplicados y los resultados oficiales por estratos para gobernador del estado de Guanajuato (elecciones extraordinarias mayo de 1995). Fuentes: Resultados oficiales desagregados por sección del IEE Guanajuato. Datos de encuesta preelectoral, encuesta de salida y conteo rápido del CEO Universidad de Guadalajara

ANEXO

A-1 Errores No Atribuibles al Muestreo

A continuación se enlistan los problemas que normalmente se presentan con más frecuencia durante la realización de encuestas de salida y conteos rápidos. Algunos son tan típicos que casi ninguna encuesta queda exenta de ellos.

- ▶ Al momento de llegar al sitio de ubicación de la casilla y cotejar el número de ésta con el asignado por la muestra, no coinciden.
- ▶ La afluencia de electores a la sección en muestra es tan baja que las cuotas en cada corte de la encuesta de salida no es posible satisfacerlas.
- ▶ Los cambios de última hora en la ubicación de la casilla que no se reportan, provocan la inexistencia de tal casilla en la dirección original.
- ▶ Algunas comunicaciones no se logran, porque no existe cobertura de telefonía de ningún tipo o bien las baterías de los teléfonos celulares se descargan.
- ▶ No es posible acceder a la casilla.
- ▶ Las secciones excesivamente aisladas arriesgan la seguridad de los encuestadores al momento de regresar.
- ▶ Tanto autoridades como militantes de partidos, cuya instrucción acerca del proceso electoral fue deficiente, impiden o entorpecen la labor de los encuestadores.
- ▶ Por diversos motivos se crean confusiones que derivan en que el encuestador trabaje en una sección equivocada.

A-2 Distribuciones de cada uno de los Ejercicios Aplicados y Municipios Completos de Guanajuato

En la tabla 1 se enlistan todos los municipios del estado de Guanajuato. Están en orden ascendente, según su población potencial con derecho a voto (mayores de 18 años según código

90 de INEGI). Además se encuentran acomodados por estratos. La columna cuyo encabezado dice "Muestreado", indica la distribución muestral que se aplicó en cada encuesta, las cuales están distinguidas con: *PR* para la encuesta preelectoral; *ES* para la encuesta de salida y *CR* para el conteo rápido. En cada una de estas columnas se ha marcado con un distintivo " ° ", si el municipio participa en la muestra correspondiente a cada ejercicio electoral.

Las encuestas y conteo muestran las respuestas en bruto que acumularon cada una de sus opciones, por lo que no van a coincidir con las respuestas válidas utilizadas a lo largo de la investigación. En el caso de la encuesta preelectoral, se mencionó en su oportunidad que el total de cuestionarios fue de 1604, pero en esta tabla, únicamente se exhiben los resultados de 1541 debido a errores de codificación.

Cve. turno	Feb 18 años	NOMBRE	Muestreado		ENCUESTA PREELECTORAL 14/Mayo/1995						ENCUESTA DE SALIDA 28/Mayo/1995						CONEJO RAPIDO 28/Mayo/1995								
			PR	ES	CR	PAN	PRI	PRD	DTRO	NING	NO SE	TOTAL	PAN	PRI	PRD	DTRO	NO RES	TOTAL	PAN	PRI	PRD	DTRO	NULO	TOTAL	
34	2023	SANTA CATARINA	*	*	*	4	4	0	0	0	1	9													
6	2561	ATARJEJA	*	*	*																				
10	4784	CORONEO	*	*	*	3	4	0	0	0	1	2	10												
36	4830	SNTGO. MARAVATIO	*	*	*																				
24	5391	PUEBLO NUEVO	*	*	*	6	1	0	0	0	0	2	9	22	5	3	0	4	34						
45	5508	XICHU	*	*	*									13	13	1	3	0	30	369	190	59	5	4	627
40	8185	TIERRA BLANCA	*	*	*	1	3	0	0	2	3	9								70	248	9	25	28	380
38	6208	TARANDACUAYO	*	*	*									28	10	2	0	14	54	202	89	15	3	9	318
13	7927	DOCTOR MORA	*	*	*	7	6	0	0	1	0	14	37	20	3	0	11	71	37	254	8	9	10	320	
22	8740	OCAMPO	*	*	*								16	7	3	4	0	30	237	181	8	5	0	431	
16	8813	IBANMARI	*	*	*	6	4	0	1	0	3	14	18	6	3	4	11	42	439	130	29	5	0	603	
43	9084	VICTORIA	*	*	*								16	8	3	2	4	33	665	532	36	12	0	1248	
	72054	TOTAL ESTRATO I	6	7	7	27	22	0	1	4	11	65	150	69	18	13	44	294	2024	1624	164	64	51	3527	
12	11254	CUTRAMARO	*	*	*	8	3	0	0	0	2	13	20	22	1	0	3	46	192	137	8	5	0	342	
29	14270	SAN DIEGO DE LA II.	*	*	*	3	3	0	0	0	3	9	19	8	1	1	0	29	217	196	13	6	0	431	
25	14306	PURISIMA DEL RCÓN.	*	*	*								17	13	0	0	2	32	127	98	3	2	7	327	
18	15719	JARAL DEL PROGRESO	*	*	*	8	3	2	0	0	5	18	17	8	1	4	0	30	323	142	81	66	8	618	
8	17807	CI. M. DOBLADO	*	*	*	8	9	1	1	0	1	20	37	17	2	4	17	77	256	202	62	24	7	551	
29	18767	TAPIMORO	*	*	*	7	8	2	0	1	2	20													
44	19589	VILLAGRAN	*	*	*															166	82	17	13	4	280
32	20465	SM. J. TURIBIANE	*	*	*	9	3	2	0	0	3	17													
26	20703	ROMITA	*	*	*	8	6	5	1	1	2	23	5	4	27	0	6	42	14	28	118	1	7	168	
4	22771	APASO EL ALTO	*	*	*	26	13	1	0	0	11	51	30	29	2	1	8	70	433	263	75	4	11	786	
41	23958	URUANIATO	*	*	*								19	10	1	0	8	38	341	213	13	1	5	573	
19	23967	REFECUARO	*	*	*	13	4	3	0	1	5	26	28	23	0	0	29	85	424	372	73	14	6	859	
21	25623	MOROLEON	*	*	*	19	4	1	0	2	2	28	5	12	4	0	5	26	359	58	32	4	0	463	
9	27014	COMONFORT	*	*	*	22	11	0	0	3	4	43	40	15	5	0	0	63	590	293	29	14	7	933	
35	27912	SANTA CRUZ DE J. R.	*	*	*								21	9	0	0	4	34	532	198	23	6	15	774	
5	29364	APASO EL GRANDE	*	*	*								45	12	3	0	6	66	810	275	100	18	17	1320	
	333489	TOTAL ESTRATO 2	11	13	14	131	67	17	2	8	40	265	303	182	47	10	89	630	4874	2564	647	178	92	8355	
1	34290	ADIASCO	*	*	*	12	9	0	0	0	6	27	36	18	16	0	5	75	147	277	502	5	11	942	
46	35302	YURIPUA	*	*	*	15	6	4	0	1	4	30	33	15	11	0	33	92	62	41	15	1	34	153	
33	36683	SACULUS DE LA PAZ	*	*	*	13	3	3	0	2	6	27	43	12	5	0	1	61	844	237	56	14	27	878	
11	38565	CORTAZAR	*	*	*	18	5	0	1	1	2	27	53	28	6	3	9	99	483	327	21	29	37	897	
30	39314	SAN FELIPE	*	*	*	11	13	2	0	2	3	31	64	21	1	0	5	91	695	305	44	5	11	1060	
31	40596	SANTO D. E. P. CON.	*	*	*	8	12	2	2	2	4	30	44	23	3	5	8	86	1175	552	116	27	28	1892	
14	47210	IGUALS HIDALGO	*	*	*	23	3	1	0	1	10	38	48	17	12	3	0	82	836	399	178	19	20	1452	
28	51531	SALVATEPERRA	*	*	*	25	8	3	1	0	4	41	82	26	6	0	21	135	1504	404	101	17	24	2050	
3	53102	ALLIPIE	*	*	*	18	8	1	2	3	6	36	60	37	2	0	2	101	1058	266	95	18	45	1482	
37	54658	SILAO	*	*	*	29	11	1	0	0	5	48	55	22	16	1	12	106	762	251	347	12	87	1453	
2	57558	ACAMBARO	*	*	*	20	7	12	1	0	3	43	49	44	3	1	20	117	663	570	147	21	26	1427	
15	59342	GUANAJUATO	*	*	*	18	11	6	1	1	4	41	54	18	12	0	8	92	1212	655	297	25	31	2120	
42	65573	VALLE DE SANTIAGO	*	*	*	24	17	10	0	1	6	58	51	38	9	7	22	127	907	725	58	65	56	1811	
23	68040	PENJAMO	*	*	*	24	17	2	2	0	4	49	737	350	117	21	180	1415	10822	5428	2023	281	423	18977	
	62442	TOTAL ESTRATO 1	14	14	13	258	128	48	9	14	67	524	118	43	8	1	44	214	1817	1026	294	49	57	3243	
27	108438	(Estr.) SALAMANCA	*	*	*	39	17	9	1	2	12	80	176	104	31	7	29	347	2492	1016	170	103	29	3810	
7	163012	(Estr.) CELAYA	*	*	*	57	33	5	4	5	18	122	184	59	23	3	53	322	2658	1244	254	84	43	4283	
17	167427	(Estr.) IRAPUATO	*	*	*	71	32	8	4	6	11	134	437	189	23	13	57	719	7223	2955	318	153	47	10596	
20	446299	(Estr.) LEON	*	*	*	199	87	13	9	9	34	351	2105	1006	1267	68	495	3941	31910	15757	3870	972	742	53191	
	1994161	GRAN TOTAL	35	38	38	782	386	100	30	50	193	1541													

Tabla 1. Relación de municipios en Guanajuato frecuencias totales de cada ejercicio por estrato Centro de Estudios de Opinión, Universidad de Guadalajara

BIBLIOGRAFIA

CASTREJÓN, CABALLERO JOSÉ LUIS (1996) *Tesina Tres Encuestas por Muestreo*. Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, U.N.A.M.

COCHRAN, WILLIAM G. (1990). *Técnicas de Muestreo*. CECOSA México, Primera Edición, Octava Reimpresión.

DURÁN, GARCÍA CARLOS Columna *Recopilaciones*, sección de sociales, periódico El Heraldo de México, 2 de octubre de 1992

ECO, UMBERTO (1992). *Cómo se hace una Tesis*. Gedisa, S.A. España, Primera edición

Diario Oficial de la Federación, *Programa Nacional de Población 1995-2000*, Secretaría de Gobernación, segunda sección, publicación 7 de agosto de 1995, México

Distritos Electorales Estatales, Instituto Electoral del Estado de Guanajuato, 1995

Distritos Electorales Federales, Guanajuato, 2ª Circunscripción del Instituto Federal Electoral, 1995

GUANAJUATO, *Perfil Sociodemográfico*, XI Censo General de Población y Vivienda, 1990, Instituto Nacional de Estadística Geografía e Informática (INEGI) (1992), México

GUANAJUATO, *Tabulados Básicos*, XI Censo General de Población y Vivienda, 1990, INEGI (1992), México

GUANAJUATO, *Panorama Sociodemográfico*, Encuesta Nacional de la Dinámica Demográfica, 1992 (ENADID) (1995), INEGI, México

SUKHATME, P.V., SUKHATME, B.V., SUKHATME, S. and ASOK, C. (1984). *Sampling Theory of Survey with Applications*. Indian Society of Agricultural Statistics and Iowa State University Press. Ames, Iowa. U.S.A., Tercera edición.

Reportes Metodológicos

- ▶ Guanajuato, Segunda Encuesta Preelectoral (18 de mayo de 1995).
- ▶ Guanajuato, Encuesta de Salida (28 de mayo de 1995).
- ▶ Guanajuato, Conteo Rápido (28 de mayo de 1995).

Centro de Estudios de Opinión de la Universidad de Guadalajara, Guadalajara, Jalisco, México.

STEPHAN, F.F. and Mc CARTHY, P.J. (1974). *Sampling Opinions. An Analysis of Survey Procedure*. Greenwood Press Publishers, Westport, Connecticut. U.S.A., Primera Edición, Primera Reimpresión

PROGRAMAS UTILIZADOS

Procesador de Texto:

Ami Pro de Lotus Development Corporation

Version 3.0 (en Español)

©Copyright 1988, 1992, Lotus Development Corporation

Hoja de Cálculo:

Microsoft Excel

Versión 5.0 (en Español)

Copyright© 1985-1994, Microsoft Corporation

Sistema Operativo Principal:

MS-DOS

Versión 6.20 (en Español)

Copyright© 1993, Microsoft Corporation

Sistema Operativo de Ambiente Gráfico:

Microsoft Windows para trabajo en grupo

Versión 3.11 (en Español)

Copyright© 1985-1993, Microsoft Corporation

Paquete Estadístico:

SPSS for Windows

Versión 5.0.1

Copyright© 1989-1992, SPSS Inc.

Manejador de Base de Datos:

D'BASE III Plus

Versión 1.0

Copyright© 1985-1987, Ashton-Tate

Bases de Datos:

Nombre: EEXMY95O

Descripción: Resultados oficiales desagregados por sección

Diseñada en: D'BASE III Plus

Proporcionado por: IEE Guanajuato

Nombre: ENCAPTI-7

Descripción: Captura de encuesta de salida

Diseñada en: EXCEL

Proporcionado por: Centro de Estudios de Opinión

Nombre: CONTEO1-7

Descripción: Captura de conteo rápido

Diseñada en: EXCEL

Proporcionado por: Centro de Estudios de Opinión

Nombre: SUTITUT

Descripción: Captura de casillas sustitutas en el conteo rápido

Diseñada en: EXCEL

Proporcionado por: Centro de Estudios de Opinión

Nombre: GUANAJUA

Descripción: Captura de encuesta preelectoral

Diseñada en: SPSS

Proporcionado por: Centro de Estudios de Opinión