



**UNIVERSIDAD NACIONAL AUTONOMA
DE MEXICO**

FACULTAD DE CIENCIAS

**CALCULO DE ESTIMADORES Y COMPARACION DE
ESTIMACIONES A PARTIR DE LA MUESTRA DEL XII CENSO
GENERAL DE POBLACION Y VIVIENDA 2000**

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 REDING BERNAL

DIRECTORA DE TESIS: DRA. GUILLERMINA ESLAVA GOMEZ



**FACULTAD DE CIENCIAS
UNAM**

2004



**FACULTAD DE CIENCIAS
SECCION ESCOLAR**



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.

**ESTA TESIS NO SALE
DE LA BIBLIOTECA**



UNIVERSIDAD NACIONAL
AUTÓNOMA DE
MÉXICO

Autorizo a la Dirección General de Bibliotecas de la UNAM a difundir en formato electrónico e impreso el contenido de mi trabajo recepcional.

NOMBRE: ARTURO REDING BERNAL
FECHA: 15/06/04
FIRMA: [Signature]

ACT. MAURICIO AGUILAR GONZÁLEZ
Jefe de la División de Estudios Profesionales de la
Facultad de Ciencias
Presente

Comunicamos a usted que hemos revisado el trabajo escrito:

Cálculo de estimadores y comparación de estimaciones a partir de la muestra del XII Censo General de Población y Vivienda 2000

realizado por Arturo Reding Bernal

con número de cuenta 09516675-1 , quien cubrió los créditos de la carrera de: Actuaría

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

Director de Tesis

Propietario Dra. Guillermina Eslava Gómez

Propietario Dr. Ignacio Méndez Ramírez

Propietario Dr. Luis Antonio Rincón Solís

Suplente Mat. Margarita Elvira Chávez Cano

Suplente Act. Francisco Sánchez Villarreal

[Handwritten signatures and initials: Jaime Alamillo, M. Eslava, M. Méndez, M. Rincón]

Consejo Departamental de Matemáticas

[Signature]
Act. Jaime Méndez Alamillo

SECRETARÍA DE EDUCACIÓN PÚBLICA
SECRETARÍA DE CULTURA

Dedicatoria

A mi MADRE

Margarita Bernal Alarcón, por brindarme su apoyo, paciencia y confianza, pero sobre todo por soportarme a lo largo de mi vida.

A mis mejores Amigos

José López Larita, Julieta Juárez Solares, Víctor Manuel Ruiz López, Roberto Mujica Palacios y Adolfo A. Márquez de la Garza, por su inigualable amistad y por compartir conmigo muchos de los mejores momentos de mi vida.

A mis amigos de Acondicionamiento Físico General

Adolfo Márquez, Roberto S. Soledad, Ruth López, Dalía Peralta, Karen Rosales, Sergio Alavez, Luis López, Lizette Villafán, Cynthia Lima, Leticia Martínez, Pedro Montes, Rosa Martínez, Sara Martínez, Diana Mena, Francisco R. Llerenas, Marcia, Miguel Ángel, Rosario, y a todos aquellos que me falten por mencionar, por compartir y disfrutar conmigo de tantas fiestas, competencias de atletismo, salidas de campo, excursiones y muchas experiencias agradables, las cuales las recordaré a lo largo de mi existencia.

A Daniel González Ramírez

Por tener la confianza de compartir conmigo parte de su difícil lucha por sobrevivir día con día y por formar parte de este trabajo al realizar el programa en lenguaje C, utilizado en esta tesis (Anexo 5).

Agradecimientos

A mi directora de tesis

Guillermina Eslava Gómez, por el tiempo, interés, comentarios, sugerencias, paciencia y a la vez exigencia, necesarias para la conclusión de este trabajo.

A mis sinodales

Ignacio Méndez Ramírez, Francisco Sánchez Villarreal, Luis Antonio Rincón Solís y Margarita E. Chávez Cano por su tiempo, sugerencias y disponibilidad de información.

A la UNAM

CCH Oriente, Facultad de Ciencias, I.I.M.A.S. y a todas aquellas Instituciones de la UNAM por la oportunidad de crecer personal y profesionalmente, pero sobre todo por el privilegio de formar parte de ella, ¡¡ORGULLOSAMENTE PUMA!!.

Arturo Reding Bernal

Índice General

Resumen	IV
Objetivo	V
Introducción	VI
Capítulo 1	
1. Estimaciones por entidad federativa	1
1.1 Desagregación de la muestra del XII Censo General de Población y Vivienda 2000	2
1.2 Factor de Expansión	3
1.3 Notación	3
1.4 Razón o proporción poblacional, caso general	5
1.5 Estimador de razón, caso general	6
1.6 Estimador de razón ignorando el diseño muestral a nivel estatal	7
1.7 Estimador de razón considerando el diseño muestral a nivel estatal	7

1.8	Estimador de razón ignorando el diseño muestral a nivel nacional	8
1.9	Estimador de razón considerando el diseño muestral a nivel nacional	9
1.10	Razón poblacional a nivel estatal y nacional	10
1.11	Cálculo de los estimadores	11
1.12	Error Cuadrático Medio, caso general	13
1.13	Varianza aproximada del estimador de razón, caso general	14
1.13.1	Estimación de la varianza ignorando el diseño muestral a nivel estatal y nacional	15
1.13.2	Estimación de la varianza considerando el diseño muestral a nivel estatal y nacional	19
1.14	Cálculos del error estándar	21
1.15	Intervalos de confianza, caso general	22
1.15.1	Intervalos de confianza por entidad federativa	23

Capítulo 2

2. Estimaciones por tamaño de localidad en cada entidad federativa	34	
2.1	Número de habitantes por tamaño de localidad en cada entidad federativa	35
2.2	Tamaño de muestra por localidad en cada entidad federativa	35
2.2.1	Tamaño de muestra por tamaño de localidad en cada entidad federativa	36
2.3	Estimador de razón ignorando el diseño muestral por tamaño de localidad en cada entidad federativa	37
2.4	Estimador de razón considerando el diseño muestral por tamaño de localidad en cada entidad federativa	37

2.5	Estimación de la varianza del estimador de razón por tamaño de localidad ignorando el diseño muestral	40
2.6	Estimación de la varianza del estimador de razón por tamaño de localidad considerando el diseño muestral	41
2.7	Intervalos de confianza por tamaño de localidad en cada entidad federativa	42
2.8	Factor de expansión por tamaño de localidad	44
	Conclusiones	61
	Bibliografía	64
	Anexos	65

Resumen

En esta tesis se compara la magnitud y precisión de los estimadores de razón considerando e ignorando el diseño de la muestra del XII Censo General de Población y Vivienda 2000. Al final de la tesis se presenta el anexo del diseño de la muestra, el cual es una versión aumentada del correspondiente al elaborado por el Instituto Nacional de Estadística, Geografía e Informática. En este diseño se tomó en cuenta un muestreo estratificado y por conglomerados en una sola etapa. Cuando se ignora el diseño, suponemos que el muestreo coincide con un muestreo aleatorio simple. La precisión de los estimadores ignorando diseño se mide respecto a la diferencia que existe entre éstos y los estimadores considerando el diseño, los cuales se suponen correctos. Adicionalmente se muestra el parámetro poblacional correspondiente al estimador, ya sea a nivel estatal o por tamaño de localidad. El parámetro a estimar es la proporción de hogares cuya vivienda cuenta con teléfono.

Las estimaciones que se calculan son puntuales y por intervalo. En las primeras, la magnitud se compara con el parámetro poblacional, en donde la estimación puntual más cercana al parámetro poblacional resulta cuando se considera el diseño de la muestra, mientras que la varianza resulta erróneamente subestimada cuando se ignora el diseño. Esto último se debe a que cuando conglomeramos se pierde precisión en la varianza.

Objetivo

Comparar la magnitud y precisión de los estimadores de razón bajo dos situaciones, en una se toma en cuenta el diseño de la muestra del XII Censo General de Población y Vivienda 2000; y en la otra se ignora el diseño muestral, lo cual se entiende como suponer un muestreo aleatorio simple. Al hacer la comparación de los estimadores se pretende conocer la importancia de realizar estimaciones adecuadamente, a partir de datos muestrales desagregados considerando el diseño muestral.

Adicionalmente, las estimaciones basadas en la información de la muestra del XII Censo General de Población y Vivienda 2000 serán comparadas con los valores poblacionales del Censo de la misma fecha. La característica poblacional a ser estimada es la proporción de hogares cuya vivienda cuenta con teléfono en México.

Introducción

La información que se utilizó en la presente tesis, fue obtenida a partir de los tabulados del censo¹ y de las bases de datos de la muestra del XII Censo General de Población y Vivienda 2000² (XII CGPV-2000), elaboradas por el Instituto Nacional de Estadística, Geografía e Informática (INEGI).

A través del sistema Contar 2000 (Sistema para la consulta de tabulados y base de datos de la muestra), el INEGI presenta tabulados concebidos como la producción del cuestionario básico del Censo, la cual consiste en tabulados por entidad federativa, con desglose municipal y un resumen general a nivel nacional. También incluye las bases de datos de la muestra del Censo, de las cuales se obtuvo la mayoría de la información utilizada en la tesis. La presentación de las bases de datos de la muestra consiste de un archivo para hogares y viviendas, otro para personas y otro para migración por cada entidad federativa. Los archivos con los que se trabajó en esta tesis fueron los de hogares y viviendas, los cuales entre las variables con las que cuentan están las correspondientes a las de los bienes con las que puede disponer la vivienda, como radio, televisión, lavadora, teléfono, etc. El sistema también incluye un apartado para la consulta de los

¹ Fuente: INEGI,
<http://www.inegi.gob.mx/est/contenidos/espanol/proyectos/censos/cpv2000/definitivos/NaI/tabulados/00vi15.pdf>

² Fuente: Contar 2000, Sistema para la consulta de tabulados y base de datos de la muestra, INEGI (Información disponible en CD, también se encuentra en el sitio web del INEGI, <http://www.inegi.gob.mx/est/contenidos/espanol/est1188.asp?c=1188>)

aspectos metodológicos del Censo, y para el caso específico de las bases de datos de la muestra, se incluye el diseño de la muestra entre otros aspectos.

Hoy en día, realizar un muestreo que nos brinde información acerca de la distribución de ciertas características, se ha convertido en algo común e incluso necesario, ya que en la mayoría de las veces no se cuenta con el tiempo ni los recursos suficientes para el levantamiento de un censo.

Para que a partir de una muestra se puedan obtener estimaciones muy cercanas a los parámetros poblacionales, es necesario que se tenga un buen diseño muestral, y esto también depende de lo que se quiera estudiar. En este caso se tomaron en cuenta algunos aspectos, como son; el esquema de muestreo, marco muestral, tamaño de muestra necesario que garantice determinada precisión de los estimadores, procedimiento de selección, entre otras cosas. En el anexo 1 se presenta la descripción del diseño muestral del XII CGPV-2000 tomado del apartado del sistema Contar 2000.

En ocasiones, al hacer el cálculo de las estimaciones, el diseño muestral no es tomado en cuenta y como consecuencia la precisión de los estimadores erróneamente disminuye. En esta tesis se analiza la importancia de que el diseño sea tomado en cuenta a través de la comparación de las estimaciones considerando e ignorando el diseño muestral. También se comparan las estimaciones con los valores poblacionales correspondientes al Censo

El caso en donde se ignora el diseño, suponemos que el muestreo coincide con un muestreo aleatorio simple (m.a.s.) por lo que en algunas ocasiones se hablará de estimaciones bajo m.a.s. en lugar de muestreo ignorando diseño muestral.

El parámetro poblacional a estimar es la proporción de hogares cuya vivienda particular³ cuenta con teléfono, se hace a nivel estatal, a nivel nacional y para cuatro agrupamientos de acuerdo al tamaño de la localidad. Generalmente entre mayor sea el tamaño de localidad mayor será la proporción de hogares que posean teléfono. NOTA: Debido a la forma en que se disponía de los datos las proporciones estimadas se refieren a la proporción de hogares cuya vivienda cuenta con teléfono, mientras que el parámetro poblacional se refiere a la proporción de viviendas que cuentan con teléfono. En el cuarto anexo se presenta la distribución del número de hogares por vivienda para cada estado y a nivel nacional. En este anexo se puede observar que a nivel nacional el 97.33% de las viviendas son habitadas por un solo hogar.

Debido a que una proporción es un caso particular de una razón, la manera de estimar al parámetro en estudio será a través de estimadores de razón, es

decir, $\hat{R} = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n x_i}$ para el caso de un m.a.s., donde \hat{R} es el estimador

de razón o de proporción de hogares cuya vivienda cuenta con teléfono, n es el tamaño de muestra, y_i vale 1 si en el hogar se cuenta con teléfono y vale 0 en caso contrario, x_i sólo toma el valor de 1 ya sea que en el hogar se cuente o no

con teléfono, por lo que $\sum_{i=1}^n x_i = n$ es el tamaño de muestra en la entidad

federativa. Los casos no especificados no son tomados en cuenta para el cálculo de los estimadores ni en el de los parámetros poblacionales.

En la tesis se realizan estimaciones puntuales y por intervalo. La forma de presentar las estimaciones es a través de tablas en las que generalmente

³ Tanto en el censo como en la muestra del XII CGPV-2000 sólo se tomó en cuenta a las viviendas particulares excluyendo a los refugios.

aparecen juntas las estimaciones de razón tomando y no en cuenta el diseño, y también la razón poblacional. Lo anterior es con el objeto de facilitar la comparación de la magnitud y precisión de los estimadores de forma directa. En los intervalos de confianza además de tablas se presentan figuras para observar de manera gráfica a los mismos.

Por último se presentan las conclusiones del trabajo, así como cinco anexos, en el primero se presenta de forma aumentada el diseño de la muestra del XII CGPV-2000 elaborado por el INEGI; en el segundo se presentan de forma gráfica la distribución del factor de expansión por entidad federativa, a nivel nacional y como manera representativa la distribución para cuatro estados por tamaño de localidad. En el tercer anexo se presenta la sintaxis que se utilizó en el software estadístico SUDAAN para la obtención de las estimaciones. En el cuarto anexo se presenta la distribución del número de hogares por vivienda para cada estado y a nivel nacional. En el quinto y último anexo se presenta el código fuente del programa que se utilizó para la preparación de los archivos antes de utilizar SUDAAN.

Capítulo 1

1. Estimaciones por entidad federativa

Como se mencionó en la introducción la presentación de las bases de datos de la muestra con las que se trabajó, se encuentra disponible en 32 archivos, uno para cada estado. Una de las ventajas de esto, es facilitar el manejo de la información, ya que cada archivo correspondiente a un estado es relativamente grande.

En este capítulo se presenta el tamaño y porcentaje de la muestra del XII Censo General de Población y Vivienda 2000 (XII CGPV-2000) desagregado a nivel estatal. Se muestran las fórmulas correspondientes de los estimadores de razón considerando e ignorando el diseño muestral así como la razón poblacional por entidad federativa y a nivel nacional. Las estimaciones se calcularon de forma puntual y sus correspondientes varianzas aproximadas. También se muestran las diferencias porcentuales entre la razón poblacional y los estimadores de razón considerando e ignorando el diseño. Por último se presentan los intervalos de confianza tanto de forma tabular como de forma gráfica.

1.1 Desagregación de la muestra del XII Censo General de Población y Vivienda 2000

En la Tabla 1.1 se presenta un resumen del tamaño y porcentaje de la muestra del XII CGPV-2000 desagregado por entidad federativa, en la cual se indica el número y porcentaje de municipios, número y porcentaje de estratos, número y porcentaje de unidades primarias de muestreo (UPM), número y porcentaje de hogares¹ y el número y porcentaje de personas a nivel estatal. También se adiciona una columna que indica el número de casos no especificados correspondientes a la variable posesión de teléfono en la vivienda.

A excepción de la columna que indica el tamaño de muestra de personas, la cual se obtuvo de los archivos correspondientes a las mismas, todas las demás columnas se obtuvieron de los archivos de hogares y viviendas para cada entidad federativa, todos ellos tomados de la muestra del XII CGPV-2000.

Al final de la Tabla 1.1 se indica el tamaño y porcentaje de muestra a nivel nacional, que es la suma correspondiente a los 32 estados de la república incluido el Distrito Federal. En este último el número de municipios equivale al número de sus delegaciones.

El total de casos no especificados es de 23,044 hogares, y por consiguiente el porcentaje de éstos a nivel nacional para la variable teléfono es de aproximadamente 1%. Estos casos no son tomados en cuenta para el cálculo de los estimadores.

¹ Tanto en el censo como en la muestra sólo se tomó en cuenta a los hogares de viviendas particulares excluyendo a los refugios.

1.2 Factor de Expansión

Cada vivienda que formó parte de la muestra, fue seleccionada de entre un conjunto de ellas, es decir, cada registro de la base de datos representa un determinado número de viviendas, este último también es conocido como el factor de expansión.

El factor de expansión corresponde al inverso de la probabilidad de inclusión de primer orden de cada vivienda poblacional, y como ya se había mencionado en el párrafo anterior, indica el número de viviendas de la población que representa cada vivienda de la muestra.

En el anexo 2 se presenta de forma gráfica la distribución del factor de expansión para cada una de la entidades federativas y para el caso en el que se realizan estimaciones por tamaño de localidad se presenta la distribución del factor de expansión para los estados de Baja California, Distrito Federal, Sonora y Yucatán para cada uno de sus correspondientes tamaños.

1.3 Notación

n_h : Total de hogares en la muestra correspondiente al estado h , $h=1,2,\dots,32$

$n = \sum_{h=1}^{32} n_h$: Tamaño total de hogares en la muestra a nivel nacional.

N_h : Total de viviendas en la población del estado h , $h=1,2,\dots,32$ obtenido a partir del XII CGPV-2000.

$N = \sum_{h=1}^{32} N_h$: Tamaño total de viviendas en la población a nivel nacional, obtenido a partir del XII CGPV-2000.

y_{hi} : Valor del elemento i -ésimo de la muestra en el estado h , $h=1,2,\dots,32$. En este caso sólo toma dos valores, 1 si en hogar se cuenta con teléfono y 0 en caso contrario.

x_{hi} : Valor del elemento i -ésimo de la muestra en el estado h , $h=1,2,\dots,32$. En este caso sólo toma el valor de 1, ya sea que en el hogar se cuente o no con teléfono, es decir, $\sum_{i=1}^{n_h} x_{hi} = n_h$.

$F_{hi} = \frac{1}{\pi_{hi}}$: Factor de expansión de la unidad i -ésima en la muestra, donde

$\pi_{hi} = P(u_{hi} \in \text{muestra})$, u_{hi} = unidad muestral i -ésima del estado h , $h=1,2,\dots,32$.

$W_h = \frac{N_h}{N}$: Peso o ponderación del estado h , $h=1,2,\dots,32$, $\sum_{h=1}^{32} W_h = 1$.

$f_h = \frac{n_h}{N_h}$: Fracción de muestreo correspondiente al estado h , $h=1,2,\dots,32$.

R_h : Proporción o razón poblacional de viviendas que cuentan con teléfono en el estado h , $h=1,2,\dots,32$.

R : Proporción o razón poblacional de viviendas que cuentan con teléfono a nivel nacional.

$\hat{R}_{m.a.s.h}$: Estimador de la razón de hogares cuya vivienda cuenta con teléfono en el estado h , $h=1,2,\dots,32$ ignorando el diseño muestral, o lo que es lo equivalente a suponer que se tomó un muestreo aleatorio simple (m.a.s.) dentro de cada estado.

$\hat{R}_{m.a.s.}$: Estimador de la razón de hogares cuya vivienda cuenta con teléfono a nivel nacional ignorando el diseño de la muestra o por m.a.s.

$\hat{R}_{estcong_h}$: Estimador de la razón de hogares cuya vivienda cuenta con teléfono para el estado h considerando el diseño de la muestra², es decir, tomando en cuenta el factor de expansión y un muestreo de conglomerados y estratos.

$\hat{R}_{estcong}$: Estimador de la razón de hogares cuya vivienda cuenta con teléfono a nivel nacional considerando el diseño de la muestra.

1.4 Razón o proporción poblacional, caso general

Como ya se mencionó anteriormente, una proporción es un caso particular de una razón y la fórmula para el cálculo de la razón poblacional es:

² Ver Diseño de la muestra (Anexo 1).

$R = \frac{Y}{X} = \frac{\bar{Y}}{\bar{X}}$, donde Y y X son el total de los valores de las variables poblacionales Y y X respectivamente. En el caso de esta tesis, lo que se está calculando es propiamente una proporción, ya que la variable X es constante y es el total de viviendas en la población (N). Por lo que

$$R = \frac{Y}{N} = \bar{Y} = \sum_{i=1}^N y_i / N.$$

1.5 Estimador de razón, caso general

El estimador de razón es una estadística que se obtiene del cociente $\hat{R} = \frac{\hat{y}}{\hat{x}}$ (Cochran, 1977, pág. 151), donde \hat{y} y \hat{x} son los estimadores de los totales poblacionales Y y X respectivamente. Para el caso de un m.a.s.

$$\hat{R} = \frac{\hat{y}}{\hat{x}} = \frac{N \sum_{i=1}^n y_i / n}{N \sum_{i=1}^n x_i / n} = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n x_i} = \frac{\sum_{i=1}^n y_i}{n}.$$

En este caso $\sum_{i=1}^n x_i$ representa tanto a los hogares que cuentan como a los que no cuentan con teléfono excluyendo a los casos no especificados, es decir

$$\sum_{i=1}^n x_i = n, \text{ donde } n \text{ es el tamaño de muestra, } x_i = 1, i = 1, 2, \dots, n.$$

1.6 Estimador de razón ignorando el diseño muestral a nivel estatal

La fórmula para el cálculo del estimador de razón por muestreo aleatorio

simple es la siguiente $\hat{R}_{m.a.s.h} = \frac{\sum_{i=1}^{n_h} y_{hi}}{\sum_{i=1}^{n_h} x_{hi}}$. En este caso, para el estado h se

tiene que $\hat{R}_{m.a.s.h} = \frac{\sum_{i=1}^{n_h} y_{hi}}{n_h}$, ya que $x_{hi} = 1, i = 1, 2, \dots, n_h$.

1.7 Estimador de razón considerando el diseño muestral a nivel estatal

Consideremos que a cada unidad última de muestreo se le asocian dos variables y_{hkji} y x_{hkji} , que correspondería al valor que toma el i -ésimo elemento $i = 1, \dots, m_j$, en la j -ésima UPM $j = 1, \dots, n_k$, en el k -ésimo estrato $k = 1, \dots, L_h$, en el estado $h, h = 1, \dots, 32$. y_{hkji} toma valores de 1 si en el hogar se cuenta con teléfono y 0 en caso contrario y x_{hkji} sólo toma el valor de 1 ya sea que en el hogar se cuente o no con teléfono y que no sea un caso no especificado. Cada uno de estos valores tiene asociado un factor de expansión F_{hkji} .

Para cada una de estas variables se estima su total por estado de la siguiente manera:

$$\hat{y}_h = \sum_{k=1}^{L_h} \sum_{j=i}^{n_k} \sum_{i=1}^{m_j} F_{hkji} y_{hkji}, \quad \hat{x}_h = \sum_{k=1}^{L_h} \sum_{j=i}^{n_k} \sum_{i=1}^{m_j} F_{hkji} x_{hkji} = \sum_{k=1}^{L_h} \sum_{j=i}^{n_k} \sum_{i=1}^{m_j} F_{hkji}$$

donde L_h es el número total de estratos en el estado h, n_k es el número total de UPM en el estrato k, m_j es el número total de hogares en la j-ésima UPM, \hat{y}_h es el estimador del número total de hogares que cuentan con teléfono en el estado h, \hat{x}_h es el estimador del número total de hogares en el estado h excluyendo a los casos no especificados. Por lo tanto la expresión para el cálculo del estimador de razón considerando el diseño muestral a nivel estatal queda de la siguiente manera:

$$\hat{R}_{estcong_h} = \frac{\hat{y}_h}{\hat{x}_h} = \frac{\sum_{k=1}^{L_h} \sum_{j=i}^{n_k} \sum_{i=1}^{m_j} F_{hkji} y_{hkji}}{\sum_{k=1}^{L_h} \sum_{j=i}^{n_k} \sum_{i=1}^{m_j} F_{hkji}} \quad (1.7.1)$$

1.8 Estimador de razón ignorando el diseño muestral a nivel nacional

El estimador de razón a nivel nacional por muestreo aleatorio simple y estratificado se obtuvo de la siguiente manera:

$$\begin{aligned}\widehat{R}_{m.a.s.} &= \sum_{h=1}^{32} \frac{N_h}{N} \widehat{R}_{m.a.s.,h} \\ &= \sum_{h=1}^{32} W_h \widehat{R}_{m.a.s.,h}\end{aligned}\quad (1.8.1)$$

donde: $W_h = \frac{N_h}{N}$ es el peso o ponderación del estado h. La ecuación (1.8.1)

se basó de Kish, L. (1979, pág. 246) donde $\widehat{R}_{m.a.s.}$ se calcula como una estimación separada de razón en la cual primero se estima la razón en cada estado y por último se pondera a nivel nacional.

Otra forma de calcular el estimador de razón ignorando el diseño muestral a nivel nacional es utilizando la siguiente ecuación (estimador combinado de razón):

$$\widehat{R}_{m.a.s.} = \frac{\sum_{h=1}^{32} \sum_{i=1}^{n_h} y_{hi}}{\sum_{h=1}^{32} \sum_{i=1}^{n_h} x_{hi}} = \frac{\sum_{h=1}^{32} \sum_{i=1}^{n_h} y_{hi}}{\sum_{h=1}^{32} n_h} \quad (1.8.2)$$

Para el cálculo del estimador de razón a nivel nacional ignorando el diseño muestral se empleó la expresión (1.8.1)

1.9 Estimador de razón considerando el diseño muestral a nivel nacional

El estimador de razón a nivel nacional por muestreo de conglomerados y estratos se calculó de la siguiente manera (estimador separado de razón):

$$\begin{aligned}\widehat{R}_{estcong} &= \sum_{h=1}^{32} \frac{N_h}{N} \widehat{R}_{estcong_h} \\ &= \sum_{h=1}^{32} W_h \widehat{R}_{estcong_h}\end{aligned}\quad (1.9.1)$$

La ecuación (1.9.1) también se basó de Kish, L (1979, pág. 246), pero a diferencia de (1.8.1) aquí los estimadores en cada estado se calcularon bajo el diseño de la muestra.

Otra manera de calcular el estimador de razón considerando el diseño muestral nivel nacional es de la siguiente forma (estimador combinado de razón):

$$\widehat{R}_{estcong} = \frac{\sum_{h=1}^{32} \sum_{k=1}^{L_h} \sum_{j=i}^{n_k} \sum_{i=1}^{m_j} F_{hkji} y_{hkji}}{\sum_{h=1}^{32} \sum_{k=1}^{L_h} \sum_{j=i}^{n_k} \sum_{i=1}^{m_j} F_{hkji}} \quad (1.9.2)$$

Para el cálculo del estimador de razón a nivel nacional considerando el diseño muestral se empleó la expresión (1.9.1)

1.10 Razón poblacional a nivel estatal y nacional

La razón o proporción poblacional para el estado h se calculó usando información desagregada a nivel estatal del XII CGPV-2000 de la siguiente manera:

$$R_h = \frac{\sum_{i=1}^{N_h} y_{hi}}{\sum_{i=1}^{N_h} x_{hi}} = \frac{\sum_{i=1}^{N_h} y_{hi}}{N_h} \quad (1.10.1)$$

A nivel nacional la razón poblacional se obtuvo a través de la siguiente fórmula:

$$R = \frac{\sum_{h=1}^{32} \sum_{i=1}^{N_h} y_{hi}}{\sum_{h=1}^{32} \sum_{i=1}^{32} x_{hi}} \quad (1.10.2)$$

$$= \frac{\sum_{h=1}^{32} \sum_{i=1}^{N_h} y_{hi}}{\sum_{h=1}^{32} N_h} = \frac{\sum_{h=1}^{32} \sum_{i=1}^{N_h} y_{hi}}{N}$$

(1.10.1) y (1.10.2) se basaron de Kish L. (1979, pág. 221)

1.11 Cálculo de los estimadores

Para el cálculo de los estimadores de razón y el de su respectiva varianza aproximada considerando e ignorando el diseño muestral para cada uno de los estados, se utilizó la sección de estimadores de razón del software estadístico especializado en muestreo (SUDAAN)³. Éste maneja varios tipos de diseño de muestra, incluyendo el utilizado en la muestra del XII CGPV-2000 (muestreo estratificado y por conglomerados).

En la Tabla 1.2 se muestran los numeradores y denominadores que serán utilizados posteriormente para el cálculo de los estimadores de razón y de la razón

³ El código que se utilizó en SUDAAN para cada uno de los cálculos de los estimadores se puede ver en el anexo 3.

poblacional. Todos los datos son utilizados a excepción de los totales a nivel nacional de las primeras 4 columnas, los cuales se incluyen en la tabla de forma puramente informativa. Los cálculos de los estimadores de razón a nivel nacional tomando y no en cuenta el diseño muestral se llevó a cabo como se indica en las ecuaciones (1.9.1) y (1.8.1) respectivamente.

En la Tabla 1.3 se presentan los estimadores de razón tomando y no en cuenta el diseño muestral, así como la razón poblacional a nivel estatal y nacional. Se observa que la estimación tomando en cuenta el diseño muestral es muy aproximada a la razón poblacional a nivel estatal y nacional con una diferencia menor a 2%, mientras que si se compara la razón censal y el estimador ignorando el diseño muestral la aproximación anterior, sólo se cumple en excepciones como en Baja California, el Distrito Federal, Quintana Roo y Tabasco.

En la Tabla 1.4 se puede observar la diferencia porcentual que existe entre la razón poblacional comparada con los estimadores de la misma y la diferencia entre estos últimos tomando y no en cuenta el diseño muestral. El promedio de las diferencias entre la razón poblacional (R_h) y el estimador en el que se considera el diseño muestral $(\hat{R}_{estcong_h})$ de los estados es aproximadamente 0.5 % y a nivel nacional la diferencia es 0.14 %. Esta última diferencia es prácticamente nula para todas las Entidades Federativas. Lo anterior no se cumple si se compara la razón poblacional (R_h) y el estimador ignorando el diseño muestral $(\hat{R}_{m.a.s-h})$ ya que a excepción de Baja California, Distrito Federal, Quintana Roo y Tabasco existe una diferencia significativa que varía de 3.31% a 14.7%, y la diferencia promedio en los estados es de aproximadamente 6.5 %.

A nivel nacional esta última diferencia es de 6.3 %. En este caso la diferencia es considerable, ya que si nos basamos en el estimador bajo m.a.s. resulta que el 30 % de los hogares de viviendas particulares en el país cuentan con teléfono, cuando en realidad el 36 % de éstas disponen del bien.

Quizá la diferencia de mayor interés, es la existente entre $\hat{R}_{m.a.s.,h}$ y $\hat{R}_{estcong,h}$, la cual es muy parecida a la distancia absoluta entre $\hat{R}_{m.a.s.,h}$ y R_h . Con lo anterior se muestra: a) No debe ignorarse el diseño muestral y b) las estimaciones considerando el diseño de la muestra son, en este caso particular, muy cercanas a su valor poblacional con una diferencia menor a 1.5%.

1.12 Error Cuadrático Medio, caso general

El error cuadrático medio de \hat{R} se define como:

$$ECM(\hat{R}) = E\left[(\hat{R} - R)^2\right] = Var(\hat{R}) + \left[E(\hat{R}) - R\right]^2 \quad (1.12.1)$$

donde $\left[E(\hat{R}) - R\right]$ es conocido como el sesgo, es decir,

$$ECM(\hat{R}) = Var(\hat{R}) + \left(\text{sesgo}(\hat{R})\right)^2$$

En general, la estimación de razón tiene un sesgo del orden $1/n$ (Cochran 1977, pág. 160). En la mayoría de las muestras en las que se realiza una estimación de razón, este sesgo es pequeño y se vuelve despreciable cuando aumenta n (el tamaño de la muestra). El sesgo tiene una cota superior, que

puede estimarse para así poder decir que el sesgo es despreciable. Al ver Cochran (1977, pág. 162) se tiene que para un m.a.s.:

$$\left| \text{sesgo}(\hat{R}) \right| \leq \sigma_{\hat{R}} \left(\frac{\sigma_{\bar{x}}}{\bar{X}} \right) = \sigma_{\hat{R}} \cdot CV(\bar{x}),$$

$$\text{donde } \sigma_{\bar{x}} = \sqrt{\text{Var}(\bar{x})} = \sqrt{\left(1 - \frac{n}{N}\right) \frac{1}{n} \cdot S_x}, \quad S_x = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{X})^2}{N-1}}$$

y $\sigma_{\hat{R}} = \sqrt{\text{Var}(\hat{R})}$. En el caso particular de esta tesis $x_i = 1$ para toda

$$i = 1, 2, \dots, N \quad \text{y} \quad \bar{X} = \frac{1}{N} \sum_{i=1}^N x_i = \frac{1}{N} N = 1 \Rightarrow S_x = 0 \quad \text{y}$$

$$\sigma_{\bar{x}} = 0 = CV(\bar{x}) \quad \therefore \left| \text{sesgo}(\hat{R}) \right| \leq 0, \text{ pero cualquier número real en}$$

valor absoluto es mayor o igual a cero, por lo tanto $\text{sesgo}(\hat{R}) = 0$ y por ende

en la tesis se limita al cálculo de la varianza. Para el caso de un diseño estratificado por conglomerados, el sesgo del estimador de razón también es cero, suponiendo que la variable del denominador vale uno en todas las observaciones.

1.13 Varianza aproximada del estimador de razón, caso general

La fórmula para el cálculo de la varianza aproximada del estimador de razón bajo m.a.s. es:

$$V(\hat{R}) \doteq \frac{(1-f)}{n\bar{X}^2} \left(\frac{\sum_{i=1}^N (y_i - Rx_i)^2}{N-1} \right) \quad (1.13.1)$$

donde $f = \frac{n}{N}$, $\bar{X} = \frac{\sum_{i=1}^N x_i}{N}$ y $R = \frac{Y}{X} = \frac{\sum_{i=1}^N y_i}{\sum_{i=1}^N x_i}$

la ecuación (1.13.1) se basó de Cochran (1977, ecuación (6.4), pág. 153)

1.13.1 Estimación de la varianza ignorando el diseño muestral a nivel estatal y nacional

La fórmula para el cálculo de la estimación de la varianza para un m.a.s. en forma general es la siguiente:

$$\hat{V}_1(\hat{R}) = \frac{(1-f)}{n\bar{X}^2} \left(\hat{S}_y^2 + \hat{R}^2 \hat{S}_x^2 - 2\hat{R}\hat{S}_{yx} \right) \quad (1.13.1.1)$$

donde $\hat{R} = \frac{\bar{y}}{\bar{x}}$, $\hat{S}_y^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1}$, $\hat{S}_x^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$ y

$$\hat{S}_{yx} = \frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{n-1}$$

Sin embargo, a veces \bar{X} no necesariamente se conoce, lo cual sugiere la siguiente forma alternativa:

$$\widehat{V}_2(\widehat{R}) = \frac{(1-f)}{n\bar{x}^2} \left(\widehat{S}_y^2 + \widehat{R}^2 \widehat{S}_x^2 - 2\widehat{R}\widehat{S}_{yx} \right) \quad (1.13.1.2)$$

Las ecuaciones (1.13.1.1) y (1.13.1.2) se basaron en Cochran (1977, ecuaciones (6.12) y (6.13) respectivamente, pág. 155). Para el cálculo de la varianza ignorando el diseño muestral en cada uno de los estados se basó en la ecuación (1.13.1.2), quedando la siguiente expresión:

$$\widehat{V}(\widehat{R}_{m.a.s.h}) = \frac{(1-f_h)}{n_h \bar{x}_h^2} \left(\widehat{S}_{y_h}^2 + \widehat{R}_{m.a.s.h}^2 \widehat{S}_{x_h}^2 - 2\widehat{R}_{m.a.s.h} \widehat{S}_{yx_h} \right)$$

donde $f_h = \frac{n_h}{N_h}$, $\bar{y}_h = \frac{\sum_{i=1}^{n_h} y_{hi}}{n_h}$, $\bar{x}_h = \frac{\sum_{i=1}^{n_h} x_{hi}}{n_h}$,

$$\widehat{S}_{y_h}^2 = \frac{\sum_{i=1}^{n_h} (y_{hi} - \bar{y}_h)^2}{n_h - 1}, \quad \widehat{S}_{x_h}^2 = \frac{\sum_{i=1}^{n_h} (x_{hi} - \bar{x}_h)^2}{n_h - 1} \quad y$$

$$\widehat{S}_{yx_h} = \frac{\sum_{i=1}^{n_h} (y_{hi} - \bar{y}_h)(x_{hi} - \bar{x}_h)}{n_h - 1}$$

En el caso particular de esta tesis $x_{hi} = 1$ para toda $i = 1, 2, \dots, n_h$,

$$h = 1, 2, \dots, 32, \text{ por lo que } \bar{x}_h = \frac{\sum_{i=1}^{n_h} 1}{n_h} = \frac{n_h}{n_h} = 1 \text{ y por ende}$$

$\hat{S}_{x_h}^2 = 0 = \hat{S}_{y_{x_h}}$, y_{hi} sólo toma los valores de 1 y 0 para toda $i = 1, 2, \dots, n_h$, $h = 1, 2, \dots, 32$ por lo que $y_{hi} = y_{hi}^2$, por lo tanto:

$$\begin{aligned}
 \hat{V}(\hat{R}_{m.a.s.h}) &= \frac{(1-f_h)}{n_h} \hat{S}_{y_h}^2 = \frac{(1-f_h)}{n_h} \left(\frac{\sum_{i=1}^{n_h} (y_{hi} - \bar{y}_h)^2}{n_h - 1} \right) \\
 &= \frac{(1-f_h)}{n_h} \left(\frac{1}{n_h - 1} \sum_{i=1}^{n_h} (y_{hi}^2 - 2\bar{y}_h y_{hi} + \bar{y}_h^2) \right) \\
 &= \frac{(1-f_h)}{n_h(n_h - 1)} \left(\sum_{i=1}^{n_h} y_{hi}^2 - 2\bar{y}_h \sum_{i=1}^{n_h} y_{hi} + n_h \bar{y}_h^2 \right) \\
 &= \frac{(1-f_h)}{n_h(n_h - 1)} (n_h \bar{y}_h - 2n_h \bar{y}_h \bar{y}_h + n_h \bar{y}_h^2) \\
 &= \frac{(1-f_h)}{n_h(n_h - 1)} (n_h \bar{y}_h - n_h \bar{y}_h^2) = \frac{(1-f_h)}{n_h(n_h - 1)} n_h \bar{y}_h (1 - \bar{y}_h) \\
 \therefore \hat{V}(\hat{R}_{m.a.s.h}) &= \frac{(1-f_h)}{n_h - 1} \bar{y}_h (1 - \bar{y}_h) \tag{1.13.1.3}
 \end{aligned}$$

Recordemos que para que la ecuación (1.13.1.3) sea válida, se debe a que y_{hi} sólo toma los valores de 1 y 0. Observación: Se tiene que al ignorar el diseño muestral se supone un m.a.s. de hogares. Si se considera el diseño se tiene un m.a.s. de conglomerados al interior de cada estrato. Entonces la expresión (1.13.1.3) es la misma en ambos casos, sólo que en el primer caso f_h es la

fracción de muestreo correspondiente a los hogares en el estado h , \bar{y}_h es la media muestral de hogares cuya vivienda cuenta con teléfono en el estado h y n_h es el tamaño de muestra de hogares en el estado h , mientras que en el segundo caso f_h es la fracción de muestreo correspondiente a los conglomerados en el estado h , \bar{y}_h es la media muestral de conglomerados en el estado h y n_h es el tamaño de muestra de conglomerados en el estado h .

Por otra parte por Corchan (1977, pág. 52) sabemos que:

$$\hat{V}(\hat{P}) = \frac{N-n}{N(n-1)} \hat{P}(1-\hat{P}) \quad (1.13.1.4)$$

donde $\hat{P} = \frac{\sum_{i=1}^n y_i}{n}$. Ahora si adoptamos (1.13.1.4) al trabajo de tesis tenemos

que: $\hat{P}_{m.a.s.h} = \frac{\sum_{i=1}^{n_h} y_{hi}}{n_h}$ que sería el estimador de la proporción de hogares

cuya vivienda cuenta con teléfono en el estado h , $h = 1, \dots, 32$.

$$\begin{aligned} \Rightarrow \hat{V}(\hat{P}_{m.a.s.h}) &= \frac{N_h - n_h}{N_h(n_h - 1)} \hat{P}_{m.a.s.h} (1 - \hat{P}_{m.a.s.h}) \\ &= \left(1 - \frac{n_h}{N_h}\right) \frac{\bar{y}(1 - \bar{y})}{(n_h - 1)} = \frac{(1 - f_h)}{(n_h - 1)} \bar{y}(1 - \bar{y}) \end{aligned} \quad (1.13.1.5)$$

ya que $\hat{P}_{m.a.s.h} = \frac{\sum_{i=1}^{n_h} y_{hi}}{n_h} = \bar{y}$ con lo que se concluye que:

$$\hat{V}(\hat{R}_{m.a.s.h}) = (1.13.1.3) = (1.13.1.5) = \hat{V}(\hat{P}_{m.a.s.h}).$$
 Como ya se

había mencionado, esto último resulta del hecho de que una proporción es un caso particular de una razón.

A nivel nacional se supuso a cada estado como un estrato y la varianza se calculó de la siguiente manera:

$$\hat{V}(\hat{R}_{m.a.s.}) = \sum_{h=1}^{32} W_h^2 \hat{V}(\hat{R}_{m.a.s.h}) \quad (1.13.1.6)$$

(1.13.1.6) se basó de Kish L. (1979, pág. 246) donde cada $\hat{R}_{m.a.s.h}$ y su respectiva varianza se calculó de manera separada e independiente en cada estado, y $W_h = N_h/N$ es el peso o ponderación del estado h, h = 1,2,...,32.

1.13.2 Estimación de la varianza considerando el diseño muestral a nivel estatal y nacional

Para el cálculo del estimador de razón considerando el diseño muestral a nivel estatal, en la sección 1.7 se presentó la siguiente fórmula:

$$\hat{R}_{estcong_h} = \frac{\hat{y}_h}{\hat{x}_h} = \frac{\sum_{k=1}^{L_h} \sum_{j=i}^{n_k} \sum_{i=1}^{m_j} F_{hkji} y_{hkji}}{\sum_{k=1}^{L_h} \sum_{j=i}^{n_k} \sum_{i=1}^{m_j} F_{hkji}} \quad (\text{razón combinada})$$

Ahora para el cálculo de la estimación de la varianza considerando el diseño muestral a nivel estatal se obtiene de la siguiente manera:

$$\widehat{V}\left(\widehat{R}_{estcong_h}\right) = \sum_{k=1}^{L_h} n_k \frac{1}{n_k - 1} \sum_{j=1}^{n_k} \left(t_{hkj}^R - \bar{t}_{hk}^R\right)^2 \quad * \quad (1.13.2.1)$$

$$\text{donde } \bar{t}_{hk}^R = \frac{\sum_{j=1}^{n_k} t_{hkj}^R}{n_k}, \quad t_{hkj}^R = \sum_{i=1}^{m_j} F_{hkji} d_{hkji}^R$$

$$\text{donde } d_{hkji}^R = \frac{\left(y_{hkji} - \widehat{R}_{estcong_h} x_{hkji}\right)}{\hat{x}_h}$$

La fórmula (1.13.2.1) es la que aparece implementada en paquetes de cómputo como Sudaan y PC-CARP, y es una aproximación a la estimación de la varianza del estimador.

A nivel nacional se supuso a cada estado como un estrato, y su varianza se calculó de la siguiente manera:

$$\widehat{V}\left(\widehat{R}_{estcong}\right) = \sum_{h=1}^{32} W_h^2 \widehat{V}\left(\widehat{R}_{estcong_h}\right)$$

donde cada $\widehat{R}_{estcong_h}$ y su respectiva varianza se calculó de manera separada e independiente en cada estado, $W_h = N_h/N$ es la ponderación del estado h , $h=1, \dots, 32$.

* Fuente: PC CARP, IV (Program Algorithms), Statistical Laboratory, Iowa State University, Ames, 1986, págs. 74-76.

1.14 Cálculos del error estándar

En la tabla 1.5 se presenta el error estándar, multiplicado por 1000, ignorando el diseño (m.a.s.) y por aproximación por series de Taylor considerando el diseño muestral así como el efecto de diseño (Deff). Para fines de cálculo, se usó el paquete sudaan y se verificó que este software también usa la fórmula (1.13.2.1) para calcular la varianza por aproximación de series de Taylor. También se presenta tomando en cuenta el diseño el error estándar, calculado por Jackknife ($E.E._{diseño}^{Jackknife}$), pero sólo de forma comparativa, ya que para el cálculo de los intervalos de confianza se utilizará el error estándar bajo Taylor (cuando se considera el diseño).

Kish (1979, pág. 232) define el efecto de diseño como:

$$Deff = \frac{\hat{V}_{diseño}(\hat{R})}{\hat{V}_{m.a.s.}(\hat{R})}$$

Para el caso de esta tesis se tiene que: $Deff = \frac{\hat{V}(\hat{R}_{estcong_h})}{\hat{V}(\hat{R}_{m.a.s._h})}$ para el estado h,

h = 1, ..., 32. En la Tabla 1.5 se puede observar que el Deff varía de 6.38 correspondiente al estado de Baja California Sur hasta 62.31 correspondiente al estado de Oaxaca. A nivel nacional el Deff es de 25.71.

En la última columna de la Tabla 1.5 se agrega el cociente del error estándar tomando en cuenta el diseño muestral bajo Taylor ($E.E._{diseño}^{Taylor}$) sobre

el error estándar ignorando el diseño de la muestra ($E.E._{m.a.s.}$), ya que en el reporte del diseño de la muestra (Anexo 1) elaborada por el INEGI se define el

$$\text{efecto de diseño como: } DEFT = \sqrt{Deff} = \frac{E.E._{diseño}}{E.E._{m.a.s.}}$$

Continuando con la Tabla 1.5, se puede observar que el valor del error estándar y por ende el de la varianza ignorando diseño (m.a.s.) es menor que la correspondiente a la del diseño de la muestra, ya sea bajo aproximaciones de series de Taylor o por Jackknife (las estimaciones de la varianza bajo estos últimos son muy aproximadas). Lo anterior se debe a que cuando estratificamos, ganamos precisión en la varianza respecto al m.a.s., pero se pierde precisión cuando conglomeramos, y en este caso domina la conglomeración.

1.15 Intervalos de confianza, caso general

Como el tamaño de muestra en cada uno de los estados es suficientemente grande, se puede aplicar la aproximación normal. Cochran (1977, pág. 156) define los límites de confianza para R de la siguiente manera:

$$\text{Límite Superior (LS)} = \hat{R} + z\sqrt{\hat{V}(\hat{R})} \quad (1.15.1)$$

$$\text{Límite Inferior (LI)} = \hat{R} - z\sqrt{\hat{V}(\hat{R})} \quad (1.15.2)$$

donde Z es el valor de la ordenada correspondiente a la distribución de la normal estándar asentado en tablas que garantiza realizar estimaciones con determinada

confianza. En este caso tomamos $Z = 1.96$ que corresponde a una confianza del 95%. Observamos que cuando el tamaño de muestra no es grande, en lugar de Z debe usarse el valor correspondiente a una distribución $t(\#UPM - \#estratos)$.

1.15.1 Intervalos de confianza por entidad federativa

Para el cálculo de los intervalos de confianza a nivel estatal y nacional, se sustituyó su respectivo estimador de razón y su error estándar en las ecuaciones (1.15.1) y (1.15.2).

En la Tabla 1.6 se muestran los límites inferior y superior de los intervalos al 95% de confianza, considerado e ignorando el diseño muestral para cada una de las entidades federativas.

El tamaño del intervalo es más grande cuando se considera el diseño que cuando se ignora. Como ya se había mencionado antes, lo anterior se debe a que su error estándar es erróneamente subestimado cuando se ignora el diseño (m.a.s.).

En la Figura 1.1 se muestran gráficamente los intervalos al 95% de confianza considerando e ignorando el diseño muestral para cada entidad federativa. A pesar de que el tamaño de los intervalos tomando en cuenta el diseño (Intervalos de color rojo) es más grande, en la mayoría de las entidades federativas los intervalos calculados ignorando diseño (Intervalos de color negro) se encuentran fuera del rango del primero, quien a su vez tiene una estimación puntual muy aproximada al valor del verdadero parámetro (ver Tabla 1.3).

Por lo tanto, tomando en cuenta el diseño, podemos decir que aproximadamente el 95% de las muestras que tomemos van a contener el valor del verdadero parámetro, cosa que no se asegura cuando se ignora el diseño. Sólo en algunos casos se interceptan los dos intervalos, que es en los estados de Baja California, Distrito Federal, Quintana Roo y Tabasco.

En la Figura 1.2 se muestran los intervalos al 95% de confianza a nivel estatal con combinaciones de los estimadores de razón y los errores estándar tomando y no en cuenta el diseño muestral. En los intervalos de color verde, se toma en cuenta la estimación puntual con diseño muestral y al error estándar ignorando diseño (m.a.s.), en este caso la estimación puntual es muy aproximada a la razón poblacional, y el tamaño del intervalo es muy pequeño debido a que se está utilizando la varianza ignorando diseño. Se menciona esta combinación ya que en la práctica en algunos estudios este caso podría llegar a suceder.

En los intervalos de color azul (Figura 1.2), se toma en cuenta la estimación puntual ignorando diseño (m.a.s.) y el error estándar tomando en cuenta el diseño, esto último es lo contrario al caso anterior en cuanto a que la estimación puntual no es muy precisa y el tamaño de los intervalos es más grande. Este caso es poco común que suceda en la práctica.

La forma más común e incorrecta es utilizar tanto la estimación de razón como la varianza sin tomar en cuenta el diseño de la muestra (Intervalos de color negro, Figura 1.2).

La forma correcta de calcular la estimación puntual y por intervalo es utilizando el diseño muestral (Intervalos de color rojo, Figura 1.2).

Tabla 1.1 Tamaño absoluto de la muestra del XII Censo General de Población y Vivienda 2000 desagregado por entidad federativa

Clave del INEGI	ENTIDAD FEDERATIVA	No. de MUNIC.	Número de ESTRATOS	No. de UPM	Número de HOGARES	Número de PERSONAS	Número de casos NO ESPEC.
01	Aguascalientes	11	324	763	19,132	87,507	252
02	Baja California	5	897	1,820	36,390	145,937	450
03	Baja C. Sur	5	270	573	9,311	36,017	89
04	Campeche	11	344	702	15,686	67,081	117
05	Coahuila	38	1,470	3,289	47,433	192,266	356
06	Colima	10	336	597	12,835	50,855	86
07	Chiapas	118	1,456	2,864	96,539	471,115	1,212
08	Chihuahua	67	1,863	5,084	72,906	286,621	816
09	Distrito Federal	16	2,347	4,501	190,499	735,046	1,434
10	Durango	39	1,058	2,287	35,716	156,952	331
11	Guanajuato	46	1,772	3,268	81,402	384,619	1,171
12	Guerrero	76	1,668	3,592	75,943	358,450	964
13	Hidalgo	84	1,106	2,248	62,016	272,802	531
14	Jalisco	124	3,177	6,733	149,135	648,635	1,257
15	México	122	3,569	6,508	252,961	1,127,611	2,567
16	Michoacán	113	2,012	4,402	105,453	476,390	999
17	Morelos	33	881	1,487	37,319	156,699	304
18	Nayarit	20	501	899	18,937	79,728	107
19	Nuevo León	51	1,676	4,377	82,346	329,835	734
20	Oaxaca	570	2,088	5,209	158,123	723,204	1,673
21	Puebla	217	2,220	4,356	133,358	627,491	1,509
22	Querétaro	18	539	1,128	28,211	127,378	498
23	Quintana Roo	8	409	810	16,440	65,259	128
24	San Luis Potosí	58	964	2,490	58,296	269,376	880
25	Sinaloa	18	1,302	2,206	41,759	180,401	356
26	Sonora	72	1,452	3,930	60,008	237,338	588
27	Tabasco	17	456	841	40,038	177,546	256
28	Tamaulipas	43	1,533	3,204	58,108	229,730	550
29	Tlaxcala	60	601	1,284	30,639	146,942	330
30	Veracruz	210	3,083	6,790	187,974	824,474	1,728
31	Yucatán	106	968	2,036	53,166	232,974	366
32	Zacatecas	57	984	2,599	43,956	192,903	405
	NACIONAL	2,443	43,326	92,877	2,312,035	10,099,182	23,044

Tabla 1.1 (Continuación) Tamaño relativo de la muestra del XII Censo General de Población y Vivienda 2000 desagregado por entidad federativa

Clave del INEGI	ENTIDAD FEDERATIVA	% del No. de MUNICIPIOS	% del No. de ESTRATOS	% del No. de UPM	% del No. de HOGARES	% del No. de PERSONAS
01	Aguascalientes	0.45	0.75	0.82	0.83	0.87
02	Baja California	0.20	2.07	1.96	1.57	1.45
03	Baja C. Sur	0.20	0.62	0.62	0.40	0.36
04	Campeche	0.45	0.79	0.76	0.68	0.66
05	Coahuila	1.56	3.39	3.54	2.05	1.90
06	Colima	0.41	0.78	0.64	0.56	0.50
07	Chiapas	4.83	3.36	3.08	4.18	4.66
08	Chihuahua	2.74	4.30	5.47	3.15	2.84
09	Distrito Federal	0.65	5.42	4.85	8.24	7.28
10	Durango	1.60	2.44	2.46	1.54	1.55
11	Guanajuato	1.88	4.09	3.52	3.52	3.81
12	Guerrero	3.11	3.85	3.87	3.28	3.55
13	Hidalgo	3.44	2.55	2.42	2.68	2.70
14	Jalisco	5.08	7.33	7.25	6.45	6.42
15	México	4.99	8.24	7.01	10.94	11.17
16	Michoacán	4.63	4.64	4.74	4.56	4.72
17	Morelos	1.35	2.03	1.60	1.61	1.55
18	Nayarit	0.82	1.16	0.97	0.82	0.79
19	Nuevo León	2.09	3.87	4.71	3.56	3.27
20	Oaxaca	23.33	4.82	5.61	6.84	7.16
21	Puebla	8.88	5.12	4.69	5.77	6.21
22	Querétaro	0.74	1.24	1.21	1.22	1.26
23	Quintana Roo	0.33	0.94	0.87	0.71	0.65
24	San Luis Potosí	2.37	2.22	2.68	2.52	2.67
25	Sinaloa	0.74	3.01	2.38	1.81	1.79
26	Sonora	2.95	3.35	4.23	2.60	2.35
27	Tabasco	0.70	1.05	0.91	1.73	1.76
28	Tamaulipas	1.76	3.54	3.45	2.51	2.27
29	Tlaxcala	2.46	1.39	1.38	1.33	1.45
30	Veracruz	8.60	7.12	7.31	8.13	8.16
31	Yucatán	4.34	2.23	2.19	2.30	2.31
32	Zacatecas	2.33	2.27	2.80	1.90	1.91
	NACIONAL	100	100	100	100	100

Tabla 1.2 Valores del numerador y denominador para el cálculo de la razón del XII Censo General de Población y Vivienda 2000 y de los estimadores de razón tomando y no en cuenta el diseño muestral

Clave del INEGI	ENTIDAD FEDERATIVA	$\sum_{i=1}^{n_h} y_{hi}$	n_h	\hat{y}_h (Con diseño)	\hat{x}_h (Con diseño)	$\sum_{i=1}^{N_h} y_{hi}^*$	N_h^*
01	Aguascalientes	6592	18880	82620	204562	82009	197614
02	Baja California	20514	35940	346443	606094	312408	549886
03	Baja C. Sur	3478	9222	46345	106479	45444	102997
04	Campeche	2820	15569	34274	162465	33196	154929
05	Coahuila	16653	47077	234004	552301	223532	533806
06	Colima	4170	12749	53437	136205	48938	123766
07	Chiapas	7021	95327	97803	822326	92062	767114
08	Chihuahua	23509	72090	325099	759065	308212	723670
09	Distrito Federal	125503	189065	1459591	2186967	1387907	2084954
10	Durango	7163	35385	98709	329002	99374	318144
11	Guanajuato	21959	80231	300138	976190	280623	907559
12	Guerrero	10235	74979	132837	670369	131166	642234
13	Hidalgo	9847	61485	100746	502762	97367	487735
14	Jalisco	57973	147878	698522	1445119	658168	1365410
15	México	85479	250394	1204036	2948970	1108010	2715746
16	Michoacán	21020	104454	238346	885658	226226	838087
17	Morelos	10642	37015	135050	372734	131068	350180
18	Nayarit	3961	18830	60855	221864	62598	217439
19	Nuevo León	39941	81612	527745	917364	505416	871094
20	Oaxaca	9088	156450	95674	755587	89597	729791
21	Puebla	20726	131849	271879	1087258	257071	1019363
22	Querétaro	8048	27713	103153	305960	98068	291755
23	Quintana Roo	4584	16312	63680	217878	61374	208371
24	San Luis Potosí	9435	57416	134172	502148	126787	484842
25	Sinaloa	13632	41403	208792	581334	207117	566986
26	Sonora	18759	59420	224726	535065	223447	521360
27	Tabasco	6496	39782	75041	423941	73904	406713
28	Tamaulipas	18600	57558	268172	683371	269751	670445
29	Tlaxcala	4868	30309	42778	200928	41822	191092
30	Veracruz	26393	186246	358468	1636817	346052	1585534
31	Yucatán	7268	52800	106340	385079	105081	368705
32	Zacatecas	6227	43551	58690	303676	58140	295400
	TOTAL	632,604	2,288,991	8,188,165	22,425,538	7,791,935	21,292,721

* <http://www.inegi.gob.mx/est/contenidos/espanol/proyectos/censos/cpv2000/definitivos/Nal/tabulados/00vi15.pdf>, INEGI.

Tabla 1.3 Estimadores de razón: ignorando el diseño $(\hat{R}_{m.a.s.h})$, considerando el diseño muestral $(\hat{R}_{estcong_h})$ y razón poblacional a nivel estatal y nacional

Clave del INEGI	ENTIDAD FEDERATIVA	$\hat{R}_{m.a.s.h}$	$\hat{R}_{estcong_h}$	R_h (poblacional) *
01	Aguascalientes	0.3492	0.4039	0.4150
02	Baja California	0.5708	0.5716	0.5681
03	Baja C. Sur	0.3771	0.4353	0.4412
04	Campeche	0.1811	0.2110	0.2143
05	Coahuila	0.3537	0.4237	0.4188
06	Colima	0.3271	0.3923	0.3954
07	Chiapas	0.0737	0.1189	0.1200
08	Chihuahua	0.3261	0.4283	0.4259
09	Distrito Federal	0.6638	0.6674	0.6657
10	Durango	0.2024	0.3000	0.3124
11	Guanajuato	0.2737	0.3075	0.3092
12	Guerrero	0.1365	0.1982	0.2042
13	Hidalgo	0.1602	0.2004	0.1996
14	Jalisco	0.3920	0.4834	0.4820
15	México	0.3414	0.4083	0.4080
16	Michoacán	0.2012	0.2691	0.2699
17	Morelos	0.2875	0.3623	0.3743
18	Nayarit	0.2104	0.2743	0.2879
19	Nuevo León	0.4894	0.5753	0.5802
20	Oaxaca	0.0581	0.1266	0.1228
21	Puebla	0.1572	0.2501	0.2522
22	Querétaro	0.2904	0.3371	0.3361
23	Quintana Roo	0.2810	0.2923	0.2945
24	San Luis Potosí	0.1643	0.2672	0.2615
25	Sinaloa	0.3293	0.3592	0.3653
26	Sonora	0.3157	0.4200	0.4286
27	Tabasco	0.1633	0.1770	0.1817
28	Tamaulipas	0.3232	0.3924	0.4023
29	Tlaxcala	0.1606	0.2129	0.2189
30	Veracruz	0.1417	0.2190	0.2183
31	Yucatán	0.1377	0.2762	0.2850
32	Zacatecas	0.1430	0.1933	0.1968
	NACIONAL	0.3027	0.3645	0.3659

* Las razones estimadas se refieren a la proporción de hogares cuya vivienda cuenta con teléfono, y el parámetro poblacional se refiere al número de viviendas que cuentan con teléfono.

Tabla 1.4 Diferencias porcentuales entre la razón poblacional y los estimadores de razón y entre estimadores de razón considerando e ignorando el diseño muestral

Clave del INEGI	ENTIDAD FEDERATIVA	Diferencia porcentual (%) entre R_h y $\hat{R}_{m.a.s.h}$	Diferencia porcentual (%) entre R_h y $\hat{R}_{estcong_h}$	Diferencia porcentual (%) entre $\hat{R}_{m.a.s.h}$ y $\hat{R}_{estcong_h}$
01	Aguascalientes	6.58	1.11	5.47
02	Baja California	0.27	0.35	0.08
03	Baja C. Sur	6.41	0.60	5.81
04	Campeche	3.31	0.33	2.98
05	Coahuila	6.50	0.49	6.99
06	Colima	6.83	0.31	6.52
07	Chiapas	4.64	0.11	4.53
08	Chihuahua	9.98	0.24	10.22
09	Distrito Federal	0.19	0.17	0.36
10	Durango	10.99	1.23	9.76
11	Guanajuato	3.55	0.17	3.38
12	Guerrero	6.77	0.61	6.17
13	Hidalgo	3.95	0.08	4.02
14	Jalisco	9.00	0.13	9.13
15	México	6.66	0.03	6.69
16	Michoacán	6.87	0.08	6.79
17	Morelos	8.68	1.20	7.48
18	Nayarit	7.75	1.36	6.39
19	Nuevo León	9.08	0.49	8.59
20	Oaxaca	6.47	0.39	6.85
21	Puebla	9.50	0.21	9.29
22	Querétaro	4.57	0.10	4.67
23	Quintana Roo	1.35	0.23	1.13
24	San Luis Potosí	9.72	0.57	10.29
25	Sinaloa	3.60	0.61	2.99
26	Sonora	11.29	0.86	10.43
27	Tabasco	1.84	0.47	1.37
28	Tamaulipas	7.92	0.99	6.93
29	Tlaxcala	5.82	0.60	5.23
30	Veracruz	7.65	0.07	7.73
31	Yucatán	14.73	0.88	13.85
32	Zacatecas	5.38	0.36	5.03
	NACIONAL	6.32	0.14	6.18

Tabla 1.5 Error estándar (E.E.) multiplicado por 1000 considerando e ignorando el diseño muestral y efecto de diseño (Deff)

Clave del INEGI	ENTIDAD FEDERATIVA	$E.E._{m.a.s.}$	$E.E._{diseño}^{Taylor}$	$E.E._{diseño}^{Jackknife}$	$\frac{\widehat{V}(\widehat{R}_{estcong_h})}{\widehat{V}(\widehat{R}_{m.a.s._h})}$	$\frac{E.E._{diseño}^{Taylor}}{E.E._{m.a.s.}}$
01	Aguascalientes	3.469	15.545	15.603	20.08	4.5
02	Baja California	2.611	14.831	14.842	32.26	5.7
03	Baja C. Sur	5.047	12.749	12.781	6.38	2.5
04	Campeche	3.087	13.175	13.200	18.21	4.3
05	Coahuila	2.204	7.647	7.660	12.04	3.5
06	Colima	4.155	16.175	16.251	15.15	3.9
07	Chiapas	0.846	5.872	5.894	48.18	6.9
08	Chihuahua	1.746	8.298	8.318	22.59	4.8
09	Distrito Federal	1.086	3.589	3.590	10.92	3.3
10	Durango	2.136	12.169	12.282	32.46	5.7
11	Guanajuato	1.574	8.853	8.907	31.64	5.6
12	Guerrero	1.254	8.071	8.124	41.42	6.4
13	Hidalgo	1.479	8.983	9.016	36.89	6.1
14	Jalisco	1.270	4.811	4.815	14.35	3.8
15	México	0.948	5.512	5.517	33.81	5.8
16	Michoacán	1.241	6.286	6.295	25.66	5.1
17	Morelos	2.353	9.076	9.091	14.88	3.9
18	Nayarit	2.970	18.930	19.147	40.62	6.4
19	Nuevo León	1.750	5.783	5.789	10.92	3.3
20	Oaxaca	0.591	4.665	4.669	62.31	7.9
21	Puebla	1.002	6.819	6.825	46.31	6.8
22	Querétaro	2.727	16.855	16.973	38.20	6.2
23	Quintana Roo	3.520	18.537	18.597	27.73	5.3
24	San Luis Potosí	1.547	11.934	12.000	59.51	7.7
25	Sinaloa	2.310	10.293	10.321	19.85	4.5
26	Sonora	1.907	8.058	8.071	17.85	4.2
27	Tabasco	1.853	12.457	12.527	45.19	6.7
28	Tamaulipas	1.949	7.456	7.468	14.63	3.8
29	Tlaxcala	2.109	8.749	8.765	17.21	4.1
30	Veracruz	0.808	5.694	5.701	49.66	7
31	Yucatán	1.499	10.499	10.517	49.06	7
32	Zacatecas	1.677	8.900	8.923	28.17	5.3
	NACIONAL	0.297	1.506	1.510	25.71	5.1

Tabla 1.6 Intervalos al 95% de confianza considerando e ignorando el diseño muestral por entidad federativa

Clave del INEGI	ENTIDAD FEDERATIVA	$\hat{R}_{m.a.s.h} - 1.96^*$ $*E.E._{m.a.s.h}$	$\hat{R}_{m.a.s.h} + 1.96^*$ $*E.E._{m.a.s.h}$	$\hat{R}_{estcong_h} - 1.96^*$ $*E.E._{estcong_h}$	$\hat{R}_{estcong_h} + 1.96^*$ $*E.E._{estcong_h}$
01	Aguascalientes	0.3424	0.3560	0.3734	0.4344
02	Baja California	0.5657	0.5759	0.5425	0.6007
03	Baja C. Sur	0.3672	0.3870	0.4103	0.4602
04	Campeche	0.1751	0.1872	0.1851	0.2368
05	Coahuila	0.3494	0.3581	0.4087	0.4387
06	Colima	0.3189	0.3352	0.3606	0.4240
07	Chiapas	0.0720	0.0753	0.1074	0.1304
08	Chihuahua	0.3227	0.3295	0.4120	0.4446
09	Distrito Federal	0.6617	0.6659	0.6604	0.6744
10	Durango	0.1982	0.2066	0.2762	0.3239
11	Guanajuato	0.2706	0.2768	0.2901	0.3248
12	Guerrero	0.1340	0.1390	0.1823	0.2140
13	Hidalgo	0.1573	0.1631	0.1828	0.2180
14	Jalisco	0.3895	0.3945	0.4739	0.4928
15	México	0.3395	0.3432	0.3975	0.4191
16	Michoacán	0.1988	0.2037	0.2568	0.2814
17	Morelos	0.2829	0.2921	0.3445	0.3801
18	Nayarit	0.2045	0.2162	0.2372	0.3114
19	Nuevo León	0.4860	0.4928	0.5639	0.5866
20	Oaxaca	0.0569	0.0592	0.1175	0.1358
21	Puebla	0.1552	0.1592	0.2367	0.2634
22	Querétaro	0.2851	0.2958	0.3041	0.3702
23	Quintana Roo	0.2741	0.2879	0.2559	0.3286
24	San Luis Potosí	0.1613	0.1674	0.2438	0.2906
25	Sinaloa	0.3247	0.3338	0.3390	0.3793
26	Sonora	0.3120	0.3194	0.4042	0.4358
27	Tabasco	0.1597	0.1669	0.1526	0.2014
28	Tamaulipas	0.3193	0.3270	0.3778	0.4070
29	Tlaxcala	0.1565	0.1647	0.1958	0.2301
30	Veracruz	0.1401	0.1433	0.2078	0.2302
31	Yucatán	0.1347	0.1406	0.2556	0.2967
32	Zacatecas	0.1397	0.1463	0.1758	0.2107
	NACIONAL	0.3022	0.3033	0.3615	0.3675

Figura 1.1 Intervalos al 95% de confianza considerando e ignorando el diseño de la muestra

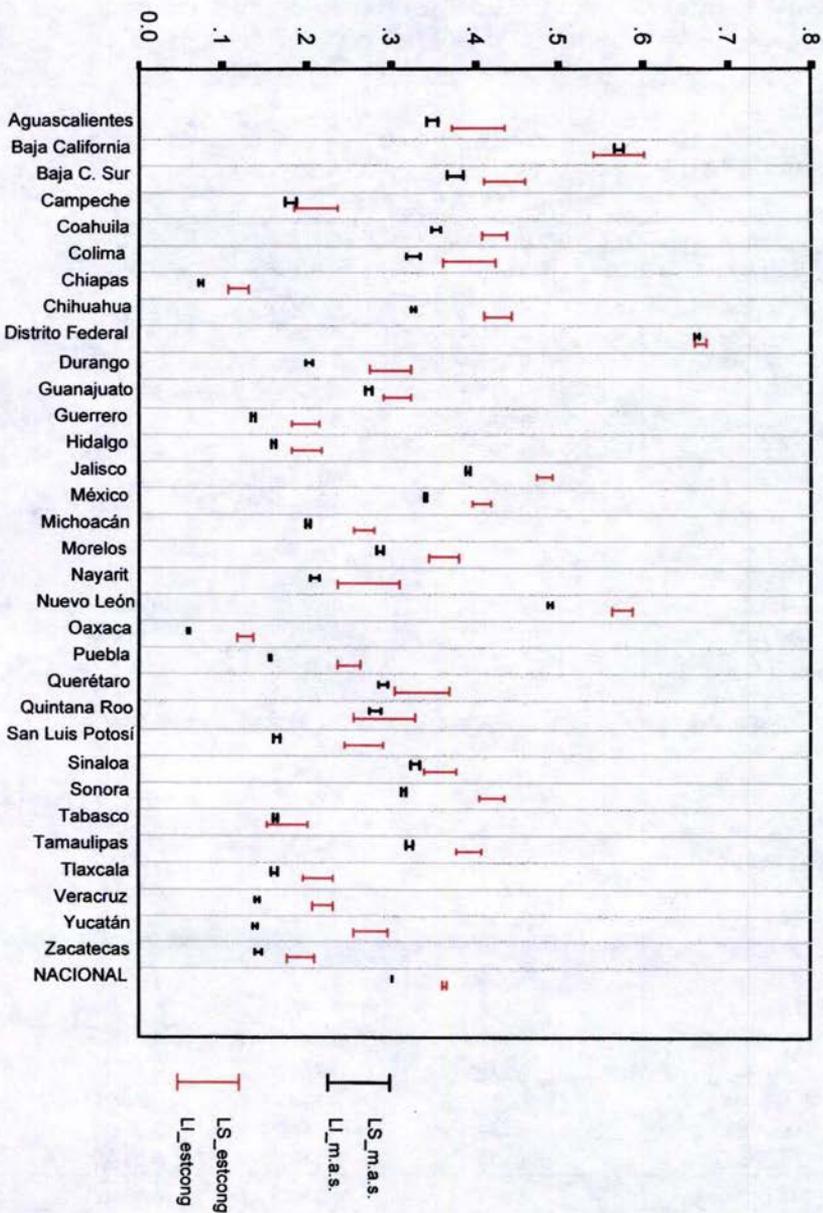
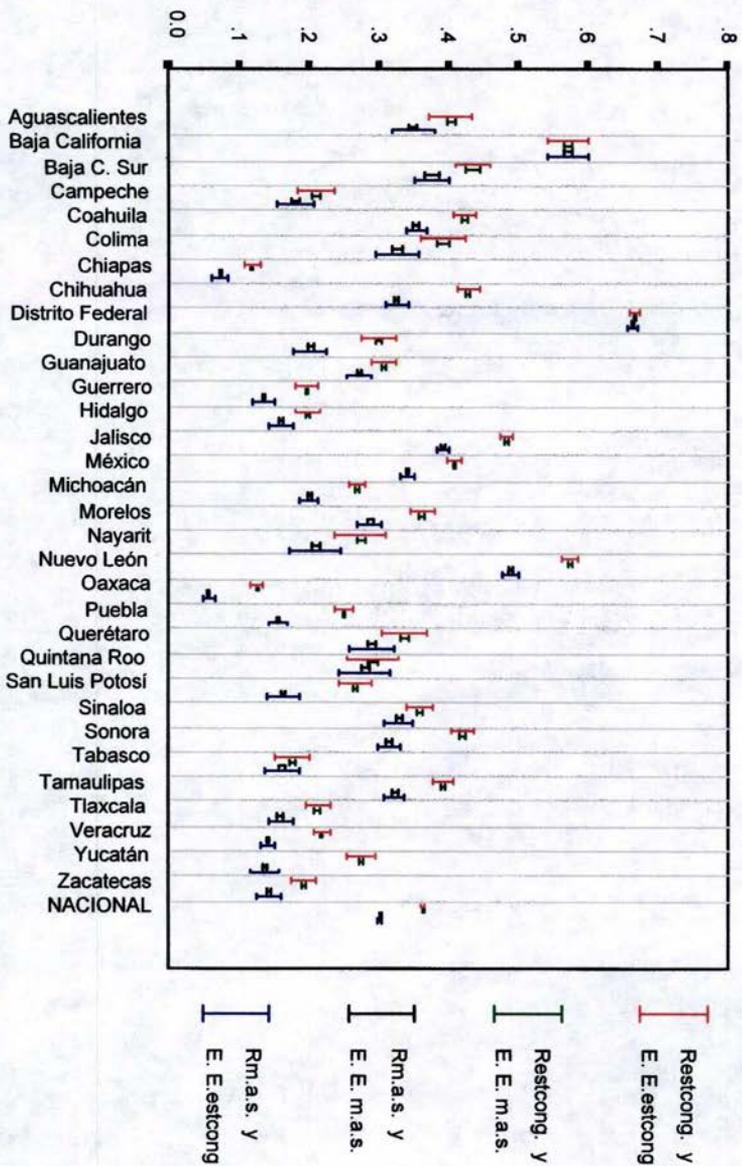


Figura 1.2 Intervalos al 95% de confianza con combinaciones de estimadores de razón ($\hat{R}_{m.d.s.}$, $\hat{R}_{esticong}$), y los errores estándar ($E.E._{m.d.s.}$ y $E.E._{esticong}$)



Capítulo 2

2. Estimaciones por tamaño de localidad en cada entidad federativa

En el capítulo 1 se calcularon estimadores a nivel estatal y nacional. En este capítulo, se realizarán las mismas estimaciones, pero a nivel de desagregación por tamaño de localidad en cada entidad federativa y nivel nacional.

El objetivo de este capítulo es estimar la proporción de hogares cuya vivienda cuenta con teléfono y su varianza aproximada de acuerdo al grado de "urbanización" de las localidades, con siete tamaños. Se presenta el número y el porcentaje de habitantes en la población, el número y porcentaje de hogares en la muestra por tamaño de localidad en cada entidad federativa y a nivel nacional. También se presenta la diferencia porcentual entre los estimadores en cada entidad federativa por tamaños de localidades. Por último se presentan los intervalos de confianza.

2.1 Número de habitantes por tamaño de localidad en cada entidad federativa

En la Tabla 2.1 se muestra el número y el porcentaje de habitantes por tamaño de localidad en cada entidad federativa basados en el XII CGPV-2000. El número de habitantes en el censo está muy relacionado con el número de hogares en muestra, ya que el porcentaje de estos últimos es muy parecido al porcentaje de habitantes del censo en la localidad correspondiente. Esta similitud se puede ver respecto a la Tabla 2.2.

2.2 Tamaño de muestra por localidad en cada entidad federativa

En la Tabla 2.2 se muestra el número y porcentaje de hogares en muestra por tamaño de localidad, así como el número total de hogares en cada entidad federativa. Para obtener el número y el respectivo porcentaje de hogares a nivel nacional, simplemente se suman los valores correspondientes a cada estado para cada tamaño de localidad.

Los datos de la Tabla 2.2 se obtuvieron de los mismos archivos de la muestra del XII CGPV-2000 de donde se obtuvo la información a nivel estatal. Esta última no está ordenada, por lo que para obtener el tamaño de muestra por tamaño de localidad, se utilizó el software estadístico SPSS. El porcentaje de la muestra (Tabla 2.2) por tamaño de localidad se calculó tomando como total el número de hogares en muestra por cada entidad federativa.

2.2.1 Tamaño de muestra por tamaño de localidad en cada entidad federativa

Los tamaños finales con los que se trabajó fueron conformados por grupos de localidades. El objetivo de formar estos últimos de acuerdo a su tamaño, es con la finalidad de hacer los cálculos de las estimaciones de razón y su respectiva varianza para cuatro grupos (formados por el grado de urbanización) y no para los siete grupos de localidades que INEGI reporta, los cuales son: las localidades que cuentan con menos de 2,500, localidades de 2,500 a 14,999, localidades de 15,000 a 19,999, localidades de 20,000 a 49,999, localidades de 50,000 a 99,999, localidades de 100,000 a 499,999 y las localidades que cuentan con 500,000 y más habitantes. El primer y el último tamaño de localidad quedan cada uno en un grupo, y son: la localidad estrictamente rural y la más urbana, por lo que las localidades que se unen en grupos son las intermedias. El tamaño de muestra por tamaño de localidad, se obtuvo sumando los hogares muestreados en las localidades correspondientes t . Los tamaños finales son:

- 1.- Localidades que cuentan con menos de 2,500 habitantes ($t = 1$).
- 2.- Localidades que cuentan de 2,500 a 49,999 habitantes ($t = 2$).
- 3.- Localidades que cuentan de 50,000 a 499,999 habitantes ($t = 3$).
- 4.- Localidades que cuentan con 500,000 y más habitantes ($t = 4$).

En la tabla 2.3 se muestra el número de hogares en muestra, el número de casos no especificados para la variable teléfono por tamaño de localidad y n_{ht} que es el número de hogares en muestra menos los casos no especificados, mismos que no son tomados en cuenta para el cálculo de los estimadores.

2.3 Estimador de razón ignorando el diseño muestral por tamaño de localidad en cada entidad federativa

El cálculo de esta razón se realizó basándonos en el capítulo 1, sección 1.6, pero en lugar de n_h se utilizó n_{ht} , que indica el número de hogares en muestra en el tamaño de localidad t , correspondiente al estado h , es decir:

$$\hat{R}_{m.a.s_h} = \frac{\sum_{i=1}^{n_{ht}} y_{hti}}{\sum_{i=1}^{n_{ht}} x_{hti}} = \frac{\sum_{i=1}^{n_{ht}} y_{hti}}{n_{ht}} \quad (2.3.1)$$

donde y_{hti} y x_{hti} son respectivamente y_{hi} y x_{hi} correspondientes a algún tamaño de localidad t .

A nivel nacional, el estimador del parámetro ignorando el diseño muestral por tamaño de localidad se calculó como un estimador combinado de razón quedando de la siguiente manera:

$$\hat{R}_{m.a.s_t} = \frac{\sum_{h=1}^{32} \sum_{i=1}^{n_{ht}} y_{hti}}{\sum_{h=1}^{32} \sum_{i=1}^{n_{ht}} x_{hti}} \quad (2.3.2)$$

2.4 Estimador de razón considerando el diseño muestral por tamaño de localidad en cada entidad federativa

El cálculo es el mismo que se utilizó en el capítulo 1, sección 1.7, ecuación (1.7.1), pero en esta ocasión por cada entidad federativa obtenemos 4 razones, una en cada tamaño de localidad, quedando de la siguiente manera:

$$\hat{R}_{estcong_h} = \frac{\hat{y}_{ht}}{\hat{x}_{ht}} = \frac{\sum_{k=1}^{L_{ht}} \sum_{j=1}^{n_{kt}} \sum_{i=1}^{m_{jt}} F_{htkji} y_{htkji}}{\sum_{k=1}^{L_{ht}} \sum_{j=1}^{n_{kt}} \sum_{i=1}^{m_{jt}} F_{htkji} x_{htkji}} = \frac{\sum_{k=1}^{L_{ht}} \sum_{j=1}^{n_{kt}} \sum_{i=1}^{m_{jt}} F_{htkji} y_{htkji}}{\sum_{k=1}^{L_{ht}} \sum_{j=1}^{n_{kt}} \sum_{i=1}^{m_{jt}} F_{htkji}} \quad (2.4.1)$$

donde L_{ht} es el número total de estratos en el tamaño de localidad t , $t = 1, \dots, 4$ correspondiente al estado h , $h = 1, \dots, 32$, n_{kt} es el número total de UPM en el estrato k , $k = 1, \dots, L_{ht}$ correspondiente al tamaño de localidad t , m_{jt} es el número total de elementos en la UPM j , $j = 1, \dots, n_{kt}$ correspondiente al tamaño de localidad t . y_{htkji} corresponde al i -ésimo hogar, j -ésima UPM, k -ésimo estrato del t -ésimo tamaño de localidad y del h -ésimo estado y toma el valor de 1 si en el hogar se cuenta con teléfono y 0 en caso contrario, x_{htkji} sólo toma el valor de 1, ya sea que en el hogar se cuente o no con teléfono y que no sea un caso no especificado. Cada una de estas dos últimas variables tiene asociado un factor de expansión F_{htkji} .

Para el cálculo del estimador considerando el diseño muestral por tamaño de localidad a nivel nacional, se obtuvo como un estimador combinado de razón, el cual se muestra en la siguiente expresión:

$$\hat{R}_{estcong_t} = \frac{\hat{y}_t}{\hat{x}_t} = \frac{\sum_{h=1}^{32} \sum_{k=1}^{L_{ht}} \sum_{j=1}^{n_{kt}} \sum_{i=1}^{m_{jt}} F_{htkji} y_{htkji}}{\sum_{h=1}^{32} \sum_{k=1}^{L_{ht}} \sum_{j=1}^{n_{kt}} \sum_{i=1}^{m_{jt}} F_{htkji} x_{htkji}} \quad (2.4.2)$$

En la Tabla 2.4 se muestran los numeradores y denominadores de los estimadores de razón considerando e ignorando el diseño de la muestra por tamaño de localidad. En esta tabla se omite el total de hogares en muestra

$\left(\sum_{i=1}^{n_{ht}} x_{hti}\right)$ debido a que ya se había incluido en la Tabla 2.3. Aquí $\sum_{i=1}^{n_{ht}} y_{hti}$ es

el total de hogares en muestra cuya vivienda cuenta con teléfono correspondientes al tamaño de localidad t , $t = 1, \dots, 4$ en el estado h , $h = 1, \dots, 32$, \hat{y}_{ht} es el total de hogares en muestra cuya vivienda cuenta con teléfono, donde cada hogar es multiplicado por su respectivo factor de expansión en el tamaño de localidad t correspondiente al estado h y \hat{x}_{ht} es la suma de los factores de expansión correspondiente a los hogares en muestra excluyendo a los casos no especificados en el t -ésimo tamaño de localidad del estado h .

En la Tabla 2.5 se muestran los estimadores puntuales de razón, así como la diferencia porcentual entre $\hat{R}_{m.a.s._{ht}}$ y $\hat{R}_{estcong_{ht}}$ en los cuatro tamaños de localidades antes mencionados. En la mayoría de los casos se puede observar que el estimador tomando en cuenta el diseño muestral es mayor que el estimador en el que se ignora. Esto último se debe a que al ignorar el diseño muestral se está subestimando al parámetro en estudio.

Se puede observar que la diferencia porcentual entre $\hat{R}_{m.a.s._{ht}}$ y $\hat{R}_{estcong_{ht}}$ varía entre 0.06 y 6.65, pero cuando se juntan cada una de éstas en el estado, la diferencia varía entre 0.08 y 13.85.

2.5 Estimación de la varianza del estimador de razón por tamaño de localidad ignorando el diseño muestral

Por el mismo argumento que se discutió en el capítulo 1, sección 1.12, el sesgo del Error Cuadrático Medio es cero, por lo que sólo se calculará la varianza de los estimadores en cada uno de los tamaños de localidad. Basándonos en la ecuación (1.13.1.2), sección 1.13.1 del capítulo 1, la fórmula utilizada para la estimación de la varianza del estimador de razón por tamaño de localidad ignorando diseño a nivel estatal queda de la siguiente manera:

$$\widehat{V}\left(\widehat{R}_{m.a.s.ht}\right) = \frac{(1-f_{ht})}{n_{ht}\bar{x}_{ht}^2} \left(\widehat{S}_{y_{ht}}^2 - \widehat{R}_{m.a.s.ht}^2 \widehat{S}_{x_{ht}}^2 - 2\widehat{R}_{m.a.s.ht} \widehat{S}_{yx_{ht}} \right) \quad (2.5.1)$$

$$\text{donde } \bar{y}_{ht} = \frac{\sum_{i=1}^{n_{ht}} y_{hti}}{n_{ht}}, \quad \bar{x}_{ht} = \frac{\sum_{i=1}^{n_{ht}} x_{hti}}{n_{ht}}, \quad \widehat{S}_{y_{ht}}^2 = \frac{\sum_{i=1}^{n_{ht}} (y_{hti} - \bar{y}_{ht})^2}{n_{ht} - 1},$$

$$\widehat{S}_{x_{ht}}^2 = \frac{\sum_{i=1}^{n_{ht}} (x_{hti} - \bar{x}_{ht})^2}{n_{ht} - 1} \quad \text{y} \quad \widehat{S}_{yx_{ht}} = \frac{\sum_{i=1}^{n_{ht}} (y_{hti} - \bar{y}_{ht})(x_{hti} - \bar{x}_{ht})}{n_{ht} - 1}.$$

A nivel nacional, el cálculo de la estimación de la varianza del estimador de razón por tamaño de localidad ignorando el diseño muestral, también se utilizó la ecuación (2.5.1), pero en este caso se toma en cuenta a los 32 estados de la república. Para lo anterior, se unió la información de los 32 estados en un solo archivo.

2.6 Estimación de la varianza del estimador de razón por tamaño de localidad considerando el diseño muestral

Para el cálculo de la estimación de la varianza del estimador de razón por tamaño de localidad considerando el diseño a nivel estatal se basó de la ecuación (1.13.2.1), sección 1.13.2 del capítulo 1 quedando de la siguiente manera:

$$\widehat{V}\left(\widehat{R}_{estcong_{ht}}\right) = \sum_{k=1}^{L_{ht}} n_{tk} \frac{1}{n_{tk} - 1} \sum_{j=1}^{n_{tk}} \left(t_{htkj}^R - \bar{t}_{htk}^R\right)^2 \quad (2.6.1)$$

$$\text{donde } \bar{t}_{htk}^R = \frac{\sum_{j=1}^{n_{tk}} t_{htkj}^R}{n_{tk}}, \quad t_{htkj}^R = \sum_{i=1}^{m_{yj}} F_{htkji} d_{htkji}^R,$$

$$d_{htkji}^R = \frac{\left(y_{htkji} - \widehat{R}_{estcong_{ht}} x_{htkji}\right)}{\widehat{x}_{ht}}$$

Para el cálculo de la estimación de la varianza del estimador de razón por tamaño de localidad considerando el diseño muestral a nivel nacional, se utiliza la ecuación (2.6.1), pero en este caso, se toma en cuenta la información de los 32 estados de la república.

En la Tabla 2.6 se muestra el error estándar multiplicado por 1000 considerando e ignorando el diseño muestral por tamaño de localidad para cada

* Fuente: PC CARP, IV (Program Algorithms), Statistical Laboratory, Iowa State University, Ames, 1986, , págs. 74-46.

una de las entidades federativas. Estos errores se calcularon por aproximación de series de Taylor. Al igual que cuando se tomaba en cuenta todo el estado, aquí también la varianza en la que se toma en cuenta el diseño muestral es mayor que cuando no se toma en cuenta. Como ya se había dicho en el capítulo anterior se debe a que cuando estratificamos, se obtiene una mayor precisión al disminuir la varianza, pero cuando conglomeramos se pierde precisión. Cuando hay una combinación de estratos y conglomerados en la varianza final se puede ganar o perder precisión, en este caso se pierde.

2.7 Intervalos de confianza por tamaño de localidad en cada entidad federativa

El tamaño de muestra en los tamaños de localidad en estudio, es suficiente para que use la aproximación normal y obtener los límites de confianza para R_{ht} , y R_t , donde las expresiones anteriores son la razón poblacional en el tamaño de localidad t del estado h , y a nivel nacional por tamaño de localidad t respectivamente. El caso general para obtener los límites al 95% de confianza se obtienen a partir de la siguiente probabilidad:

$$P\left(\hat{R} - z\sqrt{\hat{V}(\hat{R})} < R < \hat{R} + z\sqrt{\hat{V}(\hat{R})}\right) = .95 \quad (2.7.1)$$

quedando los intervalos de confianza entre los siguientes límites

$$\text{Límite Superior (LS)} = \hat{R} + z\sqrt{\hat{V}(\hat{R})} \quad (2.7.2)$$

$$\text{Límite Inferior } (LI) = \hat{R} - z\sqrt{\hat{V}(\hat{R})} \quad (2.7.3)$$

donde Z es el valor asentado en tablas de la normal estándar, para nuestro caso $z = 1.96$ que corresponde a una confianza del 95%.

Para el cálculo de los límites de confianza por tamaño de localidad en cada entidad federativa considerando e ignorando el diseño muestral se sustituyó el estimador de razón y su respectivo error estándar en las ecuaciones (2.7.2) y (2.7.3).

En la Tabla 2.7 se muestran los límites inferior y superior de los intervalos al 95% de confianza considerando e ignorando el diseño muestral por tamaño de localidades para cada una de las entidades federativas.

En la Figura 2.1 se muestran de manera gráfica los intervalos al 95% de confianza por tamaño de localidad ignorando el diseño muestral. Adicionalmente se muestran los correspondientes a la entidad federativa completa, esta última se diferencia de las demás debido a que además del intervalo, tiene indicada la estimación puntual con el símbolo de (+). Se puede observar que la proporción de hogares cuya vivienda cuenta con teléfono y por ende del intervalo de confianza, está muy relacionado con el grado de urbanización del tamaño de la localidad, es decir, la proporción de hogares cuya vivienda cuenta con teléfono en las localidades con menos de 2,500 habitantes (Intervalos de color rojo) es muy baja, en las localidades de 2,500 a 49,999 habitantes (Intervalos de color verde) es mayor que la anterior y la proporción más alta es la correspondiente a las localidades con 500,000 y más habitantes (Intervalos de color azul).

En la Figura 2.2 se muestran los intervalos al 95% de confianza por tamaño de localidad considerando el diseño muestral. Al igual que en la Figura 2.1 en ésta también se adicionan los intervalos de confianza de la entidad federativa completa. Lo anterior se realizó con la finalidad de comparar de manera gráfica los intervalos de los hogares cuya vivienda cuenta con teléfono de cada uno de los tamaños con ellos mismos y con el intervalo de la entidad federativa completa, es decir, que incluye todos los tamaños de localidad.

Si se compara la Figura 2.1 con la Figura 2.2 se puede observar que el tamaño de los intervalos cuando se considera el diseño muestral es más grande que cuando se ignora. Lo anterior se debe a lo que ya se había explicado antes respecto a la varianza.

2.8 Factor de expansión por tamaño de localidad

En las últimas 15 gráficas del anexo 2 se muestra la distribución del factor de expansión por tamaño de localidad de 4 entidades federativas que corresponden a los estados de Baja California, Distrito Federal, Sonora y Yucatán. En las dos primeras la estimación de razón de la entidad completa ignorando el diseño es muy aproximada a la correspondiente considerando diseño. Y en las dos últimas lo anterior no se cumple.

Continuando con estas últimas gráficas, se puede observar que en general el factor de expansión que representa de una a cuatro viviendas, en su mayoría caen dentro de tamaños de localidad $t= 1,2$, mientras que los factores de expansión que representan a más de 5 viviendas están distribuidos en los cuatro tamaños de localidad, y la frecuencia de ellos depende del grado de urbanización y del tamaño de muestra que se hizo de cada uno de los tamaños de localidad en cada entidad federativa.

Tabla 2.1 Número y porcentaje de habitantes por tamaño de localidad en cada entidad federativa¹

Clave del INEGI	Tamaño de localidad	Menor a 2,500 habitantes		2,500 a 14,999 habitantes		15,000 a 19,999 habitantes		20,000 a 49,999 habitantes	
	ENTIDAD FEDERATIVA	Número de Hab.	%	Número de Hab.	%	Número de Hab.	%	Número de Hab.	%
01	Aguascalientes	186706	19.8	69592	7.4	17980	1.9	75915	8.0
02	Baja California	209367	8.4	189320	7.6	65062	2.6	49178	2.0
03	Baja C. Sur	79306	18.7	77106	18.2	0	0	104675	24.7
04	Campeche	200380	29.0	124526	18.0	0	0	48946	7.1
05	Coahuila	243317	10.6	117937	5.1	90187	3.9	291049	12.7
06	Colima	78189	14.4	83737	15.4	15384	2.8	0	0
07	Chiapas	2129034	54.3	670169	17.1	52395	1.3	282123	7.2
08	Chihuahua	533460	17.5	217039	7.1	34979	1.1	89022	2.9
09	Distrito Federal	20320	0.2	80171	0.9	91265	1.1	21966	0.3
10	Durango	524606	36.2	186164	12.9	19210	1.3	22571	1.6
11	Guanajuato	1529249	32.8	411892	8.8	69641	1.5	417495	9.0
12	Guerrero	1376446	44.7	502181	16.3	68007	2.2	157513	5.1
13	Hidalgo	1132897	50.7	418684	18.7	47614	2.1	310157	13.9
14	Jalisco	976700	15.4	834219	13.2	206423	3.3	507489	8.0
15	México	1792276	13.7	1752449	13.4	341738	2.6	592418	4.5
16	Michoacán	1378901	34.6	856328	21.5	31315	0.8	393123	9.9
17	Morelos	226574	14.6	405010	26.0	98541	6.3	132704	8.5
18	Nayarit	329757	35.8	205598	22.3	51892	5.6	67121	7.3
19	Nuevo León	252770	6.6	177722	4.6	15976	0.4	148049	3.9
20	Oaxaca	1907391	55.5	758798	22.1	55187	1.6	244484	7.1
21	Puebla	1610175	31.7	1210804	23.9	97955	1.9	401186	7.9
22	Querétaro	455434	32.4	233029	16.6	15301	1.1	64596	4.6
23	Quintana Roo	153425	17.5	81362	9.3	18545	2.1	43613	5.0
24	San Luis Potosí	941729	41.0	272333	11.8	0	0	116589	5.1
25	Sinaloa	826442	32.6	374776	14.8	0	0	145560	5.7
26	Sonora	374852	16.9	249537	11.3	0	0	256163	11.6
27	Tabasco	875252	46.3	369108	19.5	49663	2.6	188323	10.0
28	Tamaulipas	401293	14.6	187379	6.8	18368	0.7	111784	4.1
29	Tlaxcala	207383	21.5	383752	39.9	35233	3.7	263065	27.3
30	Veracruz	2829007	40.9	1199121	17.4	172214	2.5	931738	13.5
31	Yucatán	309457	18.7	372937	22.5	36440	2.2	276846	16.7
32	Zacatecas	631546	46.7	267783	19.8	71951	5.3	92481	6.8
	NACIONAL	24,723,641	25.4	13,340,563	13.7	1,888,466	1.9	6,847,942	7.0

¹http://www.inegi.gob.mx/est/contenidos/espanol/proyectos/censos/cpv2000/definitivos/Nal/tabulados/00Po03_1.pdf , XII Censo General de Población y Vivienda 2000, INEGI.

Tabla 2.1 (Continuación) Número y porcentaje de habitantes por tamaño de localidad en cada entidad federativa

Clave del INEGI	Tamaño de localidad	50,000 a 99,999 habitantes		100,000 a 499,999 habitantes		500,000 y más habitantes		Entidad
	ENTIDAD FEDERATIVA	Número de Hab.	%	Número de Hab.	%	Número de Hab.	%	Población Total
01	Aguascalientes	0	0	0	0	594092	62.9	944285
02	Baja California	52394	2.1	223492	9.0	1698554	68.3	2487367
03	Baja C. Sur	0	0	162954	38.4	0	0	424041
04	Campeche	0	0	316837	45.9	0	0	690689
05	Coahuila	62930	2.7	427099	18.6	1065551	46.4	2298070
06	Colima	245678	45.3	119639	22.0	0	0	542627
07	Chiapas	70311	1.8	716860	18.3	0	0	3920892
08	Chihuahua	333256	10.9	0	0	1845151	60.4	3052907
09	Distrito Federal	0	0	3004242	34.9	5387275	62.6	8605239
10	Durango	58862	4.1	637248	44.0	0	0	1448661
11	Guanajuato	480039	10.3	733898	15.7	1020818	21.9	4663032
12	Guerrero	107341	3.5	247505	8.0	620656	20.2	3079649
13	Hidalgo	94637	4.2	231602	10.4	0	0	2235591
14	Jalisco	314914	5.0	925384	14.6	2556873	40.4	6322002
15	México	330035	2.5	3891072	29.7	4396698	33.6	13096686
16	Michoacán	427307	10.7	348697	8.7	549996	13.8	3985667
17	Morelos	85914	5.5	606553	39.0	0	0	1555296
18	Nayarit	0	0	265817	28.9	0	0	920185
19	Nuevo León	109149	2.8	1349724	35.2	1780751	46.4	3834141
20	Oaxaca	221059	6.4	251846	7.3	0	0	3438765
21	Puebla	280295	5.5	204598	4.0	1271673	25.0	5076686
22	Querétaro	99483	7.1	0	0	536463	38.2	1404306
23	Quintana Roo	59225	6.8	518793	59.3	0	0	874963
24	San Luis Potosí	64206	2.8	275295	12.0	629208	27.4	2299360
25	Sinaloa	120348	4.7	528895	20.8	540823	21.3	2536844
26	Sonora	256200	11.6	534289	24.1	545928	24.6	2216969
27	Tabasco	78637	4.2	330846	17.5	0	0	1891829
28	Tamaulipas	218777	7.9	1815621	65.9	0	0	2753222
29	Tlaxcala	73213	7.6	0	0	0	0	962646
30	Veracruz	129380	1.9	1647515	23.8	0	0	6908975
31	Yucatán	0	0	0	0	662530	40.0	1658210
32	Zacatecas	175902	13.0	113947	8.4	0	0	1353610
	NACIONAL	4,549,492	4.7	20,430,268	21.0	25,703,040	26.4	97,483,412

Tabla 2.2 Número y porcentaje de hogares en la muestra por tamaño de localidad en cada entidad federativa²

Clave del INEGI	Tamaño de localidad ENTIDAD FEDERATIVA	Menor a 2,500 habitantes		2,500 a 14,999 habitantes		15,000 a 19,999 habitantes		20,000 a 49,999 habitantes	
		Número de Hogares	%	Número de Hogares	%	Número de Hogares	%	Número de Hogares	%
01	Aguascalientes	4640	24.3	2458	12.8	536	2.8	1321	6.9
02	Baja California	2499	6.9	2326	6.4	1130	3.1	535	1.5
03	Baja C. Sur	2641	28.4	1629	17.5	0	0	2409	25.9
04	Campeche	5397	34.4	3171	20.2	0	0	918	5.9
05	Coahuila	11660	24.6	4277	9.0	2467	5.2	5061	10.7
06	Colima	3128	24.4	2167	16.9	591	4.6	0	0
07	Chiapas	58358	60.5	19926	20.6	965	1.0	4312	4.5
08	Chihuahua	31472	43.2	5918	8.1	890	1.2	1817	2.5
09	Distrito Federal	3778	2.0	1595	0.8	1902	1.0	1028	0.5
10	Durango	19728	55.2	4505	12.6	280	0.8	449	1.3
11	Guanajuato	27954	34.3	11080	13.6	1210	1.5	8304	10.2
12	Guerrero	40745	53.7	13238	17.4	1810	2.4	2510	3.3
13	Hidalgo	37122	59.9	11166	18.0	1082	1.7	6429	10.4
14	Jalisco	41496	27.8	27958	18.7	4856	3.3	10303	6.9
15	México	56760	22.4	47786	18.9	7879	3.1	11132	4.4
16	Michoacán	46981	44.6	30159	28.6	533	0.5	6647	6.3
17	Morelos	9644	25.8	11762	31.5	1942	5.2	2014	5.4
18	Nayarit	8487	44.8	4553	24.0	767	4.1	1297	6.8
19	Nuevo León	17692	21.5	8421	10.2	449	0.5	3614	4.4
20	Oaxaca	120934	76.5	24586	15.5	1049	0.7	4051	2.6
21	Puebla	65414	49.1	32598	24.4	1714	1.3	6911	5.2
22	Querétaro	10179	36.1	4647	16.5	259	0.9	1088	3.9
23	Quintana Roo	3754	22.8	2027	12.3	287	1.7	964	5.9
24	San Luis Potosí	33560	57.6	8661	14.9	0	0	1893	3.2
25	Sinaloa	13577	32.5	6431	15.4	0	0	2491	6.0
26	Sonora	23599	39.3	10755	17.9	0	0	4290	7.1
27	Tabasco	17994	44.9	9136	22.8	1309	3.3	3924	9.8
28	Tamaulipas	17372	29.9	6195	10.7	487	0.8	1920	3.3
29	Tlaxcala	10539	34.4	13627	44.5	948	3.1	3392	11.1
30	Veracruz	96826	51.5	38518	20.5	4382	2.3	18627	9.9
31	Yucatán	23292	43.8	14035	26.4	690	1.3	4742	8.9
32	Zacatecas	25893	58.9	9030	20.5	1294	2.9	2057	4.7
	NACIONAL	893,115	38.6	394,341	17.1	41,708	1.8	126,450	5.5

² Fuente: Contar 2000, Sistema para la consulta de tabulados y base de datos de la muestra, INEGI.

Tabla 2.2 (Continuación) Número y porcentaje de hogares en la muestra por tamaño de localidad en cada entidad federativa

Clave del INEGI	Tamaño de localidad	50,000 a 99,999 habitantes		100,000 a 499,999 habitantes		500,000 y más habitantes		Entidad
	ENTIDAD FEDERATIVA	Número de Hogares	%	Número de Hogares	%	Número de Hogares	%	Total de Hogares
01	Aguascalientes	0	0	0	0	10177	53.2	19132
02	Baja California	883	2.4	4377	12.0	24640	67.7	36390
03	Baja C. Sur	0	0	2632	28.3	0	0	9311
04	Campeche	0	0	6200	39.5	0	0	15686
05	Coahuila	923	1.9	7221	15.2	15824	33.4	47433
06	Colima	4908	38.2	2041	15.9	0	0	12835
07	Chiapas	1127	1.2	11851	12.3	0	0	96539
08	Chihuahua	5671	7.8	0	0	27138	37.2	72906
09	Distrito Federal	0	0	76969	40.4	105227	55.2	190499
10	Durango	949	2.7	9805	27.5	0	0	35716
11	Guanajuato	7563	9.3	11389	14.0	13902	17.1	81402
12	Guerrero	2337	3.1	4912	6.5	10391	13.7	75943
13	Hidalgo	1681	2.7	4536	7.3	0	0	62016
14	Jalisco	5897	4.0	15218	10.2	43407	29.1	149135
15	México	4846	1.9	57057	22.6	67501	26.7	252961
16	Michoacán	6707	6.4	6060	5.7	8366	7.9	105453
17	Morelos	1221	3.3	10736	28.8	0	0	37319
18	Nayarit	0	0	3833	20.2	0	0	18937
19	Nuevo León	2532	3.1	20993	25.5	28645	34.8	82346
20	Oaxaca	3716	2.4	3787	2.4	0	0	158123
21	Puebla	3805	2.9	3349	2.5	19567	14.7	133358
22	Querétaro	1989	7.1	0	0	10049	35.6	28211
23	Quintana Roo	1344	8.2	8064	49.1	0	0	16440
24	San Luis Potosí	920	1.6	3929	6.7	9333	16.0	58296
25	Sinaloa	2248	5.4	8761	21.0	8251	19.8	41759
26	Sonora	4009	6.7	9325	15.5	8030	13.4	60008
27	Tabasco	1450	3.6	6225	15.5	0	0	40038
28	Tamaulipas	2951	5.1	29183	50.2	0	0	58108
29	Tlaxcala	2133	7.0	0	0	0	0	30639
30	Veracruz	2455	1.3	27166	14.5	0	0	187974
31	Yucatán	0	0	0	0	10407	19.6	53166
32	Zacatecas	3045	6.9	2637	6.0	0	0	43956
	NACIONAL	77,310	3.3	358,256	15.5	420,855	18.2	2,312,035

Tabla 2.3 Número de hogares y casos no especificados para la variable teléfono por tamaño de localidad en muestra³

Clave del INEGI	Tamaño de localidad (<i>l</i>)	Menos de 2,500 habitantes			2,500 a 49,999 habitantes		
	ENTIDAD FEDERATIVA	Número de hogares	Casos no espec.	n_{ht}	Número de hogares	Casos no espec.	n_{ht}
01	Aguascalientes	4640	58	4582	4315	33	4282
02	Baja California	2499	41	2458	3991	51	3940
03	Baja C. Sur	2641	42	2599	4038	35	4003
04	Campeche	5397	69	5328	4089	26	4063
05	Coahuila	11660	128	11532	11805	99	11706
06	Colima	3128	28	3100	2758	12	2746
07	Chiapas	58358	910	57448	25203	205	24998
08	Chihuahua	31472	390	31082	8625	86	8539
09	Distrito Federal	3778	40	3738	4525	45	4480
10	Durango	19728	222	19506	5234	50	5184
11	Guanajuato	27954	424	27530	20594	341	20253
12	Guerrero	40745	667	40078	17558	177	17381
13	Hidalgo	37122	400	36722	18677	105	18572
14	Jalisco	41496	363	41133	43117	346	42771
15	México	56760	610	56150	66797	692	66105
16	Michoacán	46981	451	46530	37339	387	36952
17	Morelos	9644	85	9559	15718	116	15602
18	Nayarit	8487	73	8414	6617	24	6593
19	Nuevo León	17692	162	17530	12484	91	12393
20	Oaxaca	120934	1386	119548	29686	246	29440
21	Puebla	65414	862	64552	41223	397	40826
22	Querétaro	10179	224	9955	5994	113	5881
23	Quintana Roo	3754	38	3716	3278	17	3261
24	San Luis Potosí	33560	530	33030	10554	178	10376
25	Sinaloa	13577	155	13422	8922	72	8850
26	Sonora	23599	308	23291	15045	111	14934
27	Tabasco	17994	123	17871	14369	88	14281
28	Tamaulipas	17372	142	17230	8602	73	8529
29	Tlaxcala	10539	84	10455	17967	233	17734
30	Veracruz	96826	1039	95787	61527	517	61010
31	Yucatán	23292	204	23088	19467	94	19373
32	Zacatecas	25893	280	25613	12381	90	12291
	NACIONAL	893,115	10,538	882,577	562,499	5,150	557,349

³ Fuente: Contar 2000, Sistema para la consulta de tabulados y base de datos de la muestra, INEGI.

Tabla 2.3 (Continuación) Número de hogares y casos no especificados para la variable teléfono por tamaño de localidad en muestra

Clave del INEGI	Tamaño de localidad (<i>l</i>)	50,000 a 499,999 habitantes			500,000 ó más habitantes		
	ENTIDAD FEDERATIVA	Número de hogares	Casos no espec.	n_{ht}	Número de hogares	Casos no espec.	n_{ht}
01	Aguascalientes	---	---	---	10177	161	10016
02	Baja California	5260	46	5214	24640	312	24328
03	Baja C. Sur	2632	12	2620	---	---	---
04	Campeche	6200	22	6178	---	---	---
05	Coahuila	8144	37	8107	15824	92	15732
06	Colima	6949	46	6903	---	---	---
07	Chiapas	12978	97	12881	---	---	---
08	Chihuahua	5671	28	5643	27138	312	26826
09	Distrito Federal	76969	604	76365	105227	745	104482
10	Durango	10754	59	10695	---	---	---
11	Guanajuato	18952	210	18742	13902	196	13706
12	Guerrero	7249	37	7212	10391	83	10308
13	Hidalgo	6217	26	6191	---	---	---
14	Jalisco	21115	186	20929	43407	362	43045
15	México	61903	571	61332	67501	694	66807
16	Michoacán	12767	107	12660	8366	54	8312
17	Morelos	11957	103	11854	---	---	---
18	Nayarit	3833	10	3823	---	---	---
19	Nuevo León	23525	192	23333	28645	289	28356
20	Oaxaca	7503	41	7462	---	---	---
21	Puebla	7154	66	7088	19567	184	19383
22	Querétaro	1989	10	1979	10049	151	9898
23	Quintana Roo	9408	73	9335	---	---	---
24	San Luis Potosí	4849	39	4810	9333	133	9200
25	Sinaloa	11009	68	10941	8251	61	8190
26	Sonora	13334	107	13227	8030	62	7968
27	Tabasco	7675	45	7630	---	---	---
28	Tamaulipas	32134	335	31799	---	---	---
29	Tlaxcala	2133	13	2120	---	---	---
30	Veracruz	29621	172	29449	---	---	---
31	Yucatán	---	---	---	10407	68	10339
32	Zacatecas	5682	35	5647	---	---	---
	NACIONAL	435,566	3,397	432,169	420,855	3,959	416,896

Tabla 2.4 Numeradores y denominador del estimador de razón considerando e ignorando el diseño muestral por tamaño de localidad

Clave del INEGI	Tamaño de localidad t	Menos de 2,500 habitantes			2,500 a 49,999 habitantes		
	ENTIDAD FEDERATIVA	$\sum_{i=1}^{n_{ht}} y_{hti}$	\hat{y}_{ht}	\hat{x}_{ht}	$\sum_{i=1}^{n_{ht}} y_{hti}$	\hat{y}_{ht}	\hat{x}_{ht}
01	Aguascalientes	544	5065	36303	901	7470	32940
02	Baja California	595	15825	61712	1422	23425	61053
03	Baja C. Sur	381	3582	20624	1640	18964	44933
04	Campeche	60	622	43351	405	4206	38942
05	Coahuila	1212	3437	58921	3751	40626	119967
06	Colima	290	2005	18989	660	6807	24324
07	Chiapas	477	3942	392336	2284	28111	238310
08	Chihuahua	3708	14115	134273	2250	23305	83055
09	Distrito Federal	551	735	4624	1297	14341	44068
10	Durango	912	5931	114257	1052	12477	51863
11	Guanajuato	1918	23277	303629	5422	51518	188217
12	Guerrero	798	7125	275572	2987	32537	158581
13	Hidalgo	2750	14698	242231	4299	47754	179288
14	Jalisco	6022	39144	217018	13894	126585	344667
15	México	4670	25939	367417	14666	139637	566904
16	Michoacán	3591	20806	291498	8174	74068	283712
17	Morelos	1010	6746	53604	3618	41422	146882
18	Nayarit	383	3807	75710	1785	24925	80798
19	Nuevo León	3149	10913	65514	4364	33893	86062
20	Oaxaca	1894	6929	397615	4517	46230	242199
21	Puebla	2034	8476	322546	6282	63373	347197
22	Querétaro	584	3936	82286	1183	16322	72112
23	Quintana Roo	285	2430	34803	562	5729	34113
24	San Luis Potosí	906	4306	190070	1837	19442	86011
25	Sinaloa	1295	20467	184577	2146	31303	115738
26	Sonora	3775	12121	90026	4439	43698	120132
27	Tabasco	666	4807	183808	2556	25782	138235
28	Tamaulipas	1315	7967	99147	2506	24924	78731
29	Tlaxcala	841	2446	35828	3202	27870	135543
30	Veracruz	2978	22629	616186	10762	126247	552163
31	Yucatán	441	1236	66985	1536	18671	152782
32	Zacatecas	1306	6325	139252	2452	22047	96390
	NACIONAL	51,341	311,789	5,220,712	118,851	1,223,709	4,945,912

Tabla 2.4 (Continuación) Numeradores y denominador del estimador de razón considerando e ignorando el diseño muestral por tamaño de localidad

	Tamaño de localidad l	50,000 a 499,999 habitantes			500,000 ó más habitantes		
Clave del INEGI	ENTIDAD FEDERATIVA	$\sum_{i=1}^{n_h} y_{hti}$	\hat{y}_{ht}	\hat{x}_{ht}	$\sum_{i=1}^{n_h} y_{hti}$	\hat{y}_{ht}	\hat{x}_{ht}
01	Aguascalientes	---	---	---	5147	70085	135319
02	Baja California	3109	43241	69522	15388	263952	413807
03	Baja C. Sur	1457	23799	40922	---	---	---
04	Campeche	2355	29446	80172	---	---	---
05	Coahuila	3422	53335	121083	8268	136606	252330
06	Colima	3220	44625	92892	---	---	---
07	Chiapas	4260	65750	191680	---	---	---
08	Chihuahua	2742	42042	85382	14809	245637	456355
09	Distrito Federal	52833	540348	781692	70822	904167	1356583
10	Durango	5199	80301	162882	---	---	---
11	Guanajuato	8066	123061	273785	6553	102282	210559
12	Guerrero	2642	33860	86028	3808	59315	150188
13	Hidalgo	2798	38294	81243	---	---	---
14	Jalisco	9503	129088	278497	28554	403705	604937
15	México	28923	461180	976506	37220	577280	1038143
16	Michoacán	4939	73787	178867	4316	69685	131581
17	Morelos	6014	86882	172248	---	---	---
18	Nayarit	1793	32123	65356	---	---	---
19	Nuevo León	13824	206961	343127	18604	275978	422661
20	Oaxaca	2677	42515	115773	---	---	---
21	Puebla	2354	39588	110401	10056	160442	307114
22	Querétaro	850	10517	24075	5431	72378	127487
23	Quintana Roo	3737	55521	148962	---	---	---
24	San Luis Potosí	1647	28168	77113	5045	82256	148954
25	Sinaloa	5521	84119	157801	4670	72903	123218
26	Sonora	6228	95079	191955	4317	73828	132952
27	Tabasco	3274	44452	101898	---	---	---
28	Tamaulipas	14779	235281	505493	---	---	---
29	Tlaxcala	825	12462	29557	---	---	---
30	Veracruz	12653	209592	468468	---	---	---
31	Yucatán	---	---	---	5291	86433	165312
32	Zacatecas	2469	30318	68034	---	---	---
	NACIONAL	214,113	2,995,735	6,081,414	248,299	3,656,932	6,177,500

Tabla 2.5 Estimadores de razón considerando e ignorando el diseño muestral y diferencia porcentual por tamaño de localidad

Clave del INEGI	Tamaño de localidad t	Menos de 2,500 habitantes			2,500 a 49,999 habitantes		
		$\hat{R}_{m.a.s.\cdot ht}$	$\hat{R}_{estcong_{ht}}$	$\hat{R}_{estcong_{ht}}$	$\hat{R}_{m.a.s.\cdot ht}$	$\hat{R}_{estcong_{ht}}$	$\hat{R}_{estcong_{ht}}$
				$\left \hat{R}_{m.a.s.\cdot ht} \right \cdot 100$			$\left \hat{R}_{m.a.s.\cdot ht} \right \cdot 100$
01	Aguascalientes	0.1187	0.1395	2.08	0.2104	0.2268	1.64
02	Baja California	0.2421	0.2564	1.44	0.3609	0.3837	2.28
03	Baja C. Sur	0.1466	0.1737	2.71	0.4097	0.4221	1.24
04	Campeche	0.0113	0.0143	0.31	0.0997	0.1080	0.83
05	Coahuila	0.1051	0.0583	4.68	0.3204	0.3386	1.82
06	Colima	0.0935	0.1056	1.20	0.2403	0.2798	3.95
07	Chiapas	0.0083	0.0100	0.17	0.0914	0.1180	2.66
08	Chihuahua	0.1193	0.1051	1.42	0.2635	0.2806	1.71
09	Distrito Federal	0.1474	0.1590	1.15	0.2895	0.3254	3.59
10	Durango	0.0468	0.0519	0.52	0.2029	0.2406	3.76
11	Guanajuato	0.0697	0.0767	0.70	0.2677	0.2737	0.60
12	Guerrero	0.0199	0.0259	0.59	0.1719	0.2052	3.33
13	Hidalgo	0.0749	0.0607	1.42	0.2315	0.2664	3.49
14	Jalisco	0.1464	0.1804	3.40	0.3248	0.3673	4.24
15	México	0.0832	0.0706	1.26	0.2219	0.2463	2.45
16	Michoacán	0.0772	0.0714	0.58	0.2212	0.2611	3.99
17	Morelos	0.1057	0.1258	2.02	0.2319	0.2820	5.01
18	Nayarit	0.0455	0.0503	0.48	0.2707	0.3085	3.77
19	Nuevo León	0.1796	0.1666	1.31	0.3521	0.3938	4.17
20	Oaxaca	0.0158	0.0174	0.16	0.1534	0.1909	3.74
21	Puebla	0.0315	0.0263	0.52	0.1539	0.1825	2.87
22	Querétaro	0.0587	0.0478	1.08	0.2012	0.2263	2.52
23	Quintana Roo	0.0767	0.0698	0.69	0.1723	0.1679	0.44
24	San Luis Potosí	0.0274	0.0227	0.48	0.1770	0.2260	4.90
25	Sinaloa	0.0965	0.1109	1.44	0.2425	0.2705	2.80
26	Sonora	0.1621	0.1346	2.74	0.2972	0.3637	6.65
27	Tabasco	0.0373	0.0262	1.11	0.1790	0.1865	0.75
28	Tamaulipas	0.0763	0.0804	0.40	0.2938	0.3166	2.28
29	Tlaxcala	0.0804	0.0683	1.22	0.1806	0.2056	2.51
30	Veracruz	0.0311	0.0367	0.56	0.1764	0.2286	5.22
31	Yucatán	0.0191	0.0185	0.06	0.0793	0.1222	4.29
32	Zacatecas	0.0510	0.0454	0.56	0.1995	0.2287	2.92
	NACIONAL	0.0582	0.0597	0.15	0.2132	0.2474	3.42

Tabla 2.5 (Continuación) Estimadores de razón considerando e ignorando el diseño muestral y deferencia porcentual por tamaño de localidad

Clave del INEGI	Tamaño de localidad f	50,000 a 499,999 habitantes			500,000 ó más habitantes		
		$\hat{R}_{m.a.s.h}$	$\hat{R}_{estcong_h}$	$\frac{\hat{R}_{estcong_h}}{\hat{R}_{m.a.s.h}} \Big _{*}$	$\hat{R}_{m.a.s.h}$	$\hat{R}_{estcong_h}$	$\frac{\hat{R}_{estcong_h}}{\hat{R}_{m.a.s.h}} \Big _{*}$
	ENTIDAD FEDERATIVA			100			100
01	Aguascalientes	---	---	---	0.5139	0.5179	0.40
02	Baja California	0.5963	0.6220	2.57	0.6325	0.6379	0.53
03	Baja C. Sur	0.5561	0.5816	2.55	---	---	---
04	Campeche	0.3812	0.3673	1.39	---	---	---
05	Coahuila	0.4221	0.4405	1.84	0.5256	0.5414	1.58
06	Colima	0.4665	0.4804	1.39	---	---	---
07	Chiapas	0.3307	0.3430	1.23	---	---	---
08	Chihuahua	0.4859	0.4924	0.65	0.5520	0.5383	1.38
09	Distrito Federal	0.6918	0.6913	0.06	0.6778	0.6665	1.13
10	Durango	0.4861	0.4930	0.69	---	---	---
11	Guanajuato	0.4304	0.4495	1.91	0.4781	0.4858	0.77
12	Guerrero	0.3663	0.3936	2.73	0.3694	0.3949	2.55
13	Hidalgo	0.4519	0.4714	1.94	---	---	---
14	Jalisco	0.4541	0.4635	0.95	0.6634	0.6674	0.40
15	México	0.4716	0.4723	0.07	0.5571	0.5561	0.11
16	Michoacán	0.3901	0.4125	2.24	0.5192	0.5296	1.03
17	Morelos	0.5073	0.5044	0.29	---	---	---
18	Nayarit	0.4690	0.4915	2.25	---	---	---
19	Nuevo León	0.5925	0.6032	1.07	0.6561	0.6530	0.31
20	Oaxaca	0.3588	0.3672	0.85	---	---	---
21	Puebla	0.3321	0.3586	2.65	0.5188	0.5224	0.36
22	Querétaro	0.4295	0.4368	0.73	0.5487	0.5677	1.90
23	Quintana Roo	0.4003	0.3727	2.76	---	---	---
24	San Luis Potosí	0.3424	0.3653	2.29	0.5484	0.5522	0.39
25	Sinaloa	0.5046	0.5331	2.85	0.5702	0.5917	2.15
26	Sonora	0.4709	0.4953	2.45	0.5418	0.5553	1.35
27	Tabasco	0.4291	0.4362	0.71	---	---	---
28	Tamaulipas	0.4648	0.4654	0.07	---	---	---
29	Tlaxcala	0.3892	0.4216	3.25	---	---	---
30	Veracruz	0.4297	0.4474	1.77	---	---	---
31	Yucatán	---	---	---	0.5118	0.5228	1.11
32	Zacatecas	0.4372	0.4456	0.84	---	---	---
	NACIONAL	0.4954	0.4926	0.28	0.5956	0.5920	0.36

Tabla 2.6 Error estándar multiplicado por 1000 considerando e ignorando el diseño muestral por tamaño de localidad en cada entidad federativa

	Tamaño de localidad t	Menos de 2,500 habitantes		2,500 a 49,999 habitantes	
Clave del INEGI	ENTIDAD FEDERATIVA	$\sqrt{\hat{V}(\hat{R}_{m.a.s.,h})}$	$\sqrt{\hat{V}(\hat{R}_{estcong,h})}$	$\sqrt{\hat{V}(\hat{R}_{m.a.s.,h})}$	$\sqrt{\hat{V}(\hat{R}_{estcong,h})}$
01	Aguascalientes	4.779	38.020	6.230	11.408
02	Baja California	8.641	45.043	7.652	14.827
03	Baja C. Sur	6.939	41.516	7.774	13.862
04	Campeche	1.446	2.681	4.700	10.115
05	Coahuila	2.856	6.532	4.313	9.733
06	Colima	5.231	22.853	8.156	22.756
07	Chiapas	0.379	1.379	1.822	9.231
08	Chihuahua	1.839	13.206	4.768	9.891
09	Distrito Federal	5.799	32.681	6.777	22.276
10	Durango	1.512	6.870	5.586	11.398
11	Guanajuato	1.534	12.558	3.111	11.025
12	Guerrero	0.698	4.815	2.862	7.908
13	Hidalgo	1.374	6.967	3.095	9.358
14	Jalisco	1.743	17.428	2.264	5.718
15	México	1.165	6.001	1.616	5.268
16	Michoacán	1.237	6.827	2.159	6.831
17	Morelos	3.144	15.611	3.379	10.432
18	Nayarit	2.273	11.204	5.473	11.578
19	Nuevo León	2.899	14.523	4.291	11.168
20	Oaxaca	0.361	1.712	2.101	6.039
21	Puebla	0.688	2.340	1.786	6.103
22	Querétaro	2.355	7.153	5.228	33.404
23	Quintana Roo	4.366	54.161	6.615	15.250
24	San Luis Potosí	0.899	3.630	3.747	12.497
25	Sinaloa	2.549	8.882	4.556	8.818
26	Sonora	2.415	13.675	3.740	7.463
27	Tabasco	1.417	4.181	3.208	9.039
28	Tamaulipas	2.023	8.533	4.933	10.428
29	Tlaxcala	2.660	12.464	2.889	8.293
30	Veracruz	0.561	7.885	1.543	5.076
31	Yucatán	0.901	2.335	1.941	8.429
32	Zacatecas	1.375	7.610	3.605	8.555
	NACIONAL	0.259	2.099	0.549	1.886

Tabla 2.6 (Continuación) Error estándar multiplicado por 1000 considerando e ignorando el diseño muestral por tamaño de localidad en cada entidad federativa

Clave del INEGI	Tamaño de localidad t ENTIDAD FEDERATIVA	50,000 a 499,999 habitantes		500,000 ó más habitantes	
		$\sqrt{\widehat{V}(\widehat{R}_{m.a.s.m})}$	$\sqrt{\widehat{V}(\widehat{R}_{estcong,m})}$	$\sqrt{\widehat{V}(\widehat{R}_{m.a.s.m})}$	$\sqrt{\widehat{V}(\widehat{R}_{estcong,m})}$
01	Aguascalientes	---	---	4.994	10.875
02	Baja California	6.795	13.460	3.091	7.008
03	Baja C. Sur	9.708	16.763	---	---
04	Campeche	6.180	12.544	---	---
05	Coahuila	5.486	10.268	3.981	8.712
06	Colima	6.005	16.789	---	---
07	Chiapas	4.145	13.249	---	---
08	Chihuahua	6.654	10.773	3.036	6.311
09	Distrito Federal	1.671	6.203	1.446	4.381
10	Durango	4.833	10.123	---	---
11	Guanajuato	3.617	11.583	4.267	8.794
12	Guerrero	5.674	10.652	4.754	13.871
13	Hidalgo	6.326	21.661	---	---
14	Jalisco	3.442	12.355	2.278	5.423
15	México	2.016	6.863	1.922	7.692
16	Michoacán	4.335	11.850	5.481	12.289
17	Morelos	4.592	14.946	---	---
18	Nayarit	8.072	22.344	---	---
19	Nuevo León	3.217	7.198	2.821	5.888
20	Oaxaca	5.553	18.597	---	---
21	Puebla	5.595	13.704	3.589	12.838
22	Querétaro	11.130	23.861	5.002	13.506
23	Quintana Roo	5.071	20.325	---	---
24	San Luis Potosí	6.843	17.715	5.189	12.018
25	Sinaloa	4.780	11.662	5.471	9.420
26	Sonora	4.340	8.512	5.582	9.683
27	Tabasco	5.667	15.809	---	---
28	Tamaulipas	2.797	5.854	---	---
29	Tlaxcala	10.592	27.517	---	---
30	Veracruz	2.885	9.042	---	---
31	Yucatán	---	---	4.916	16.564
32	Zacatecas	6.602	17.925	---	---
	NACIONAL	0.761	2.382	0.76	2.227

Tabla 2.7 Intervalos al 95% de confianza considerando e ignorando el diseño muestral por tamaño de localidad en cada entidad federativa

Clave del INEGI	Tamaño de localidad t ENTIDAD FEDERATIVA	Menores a 2,500 habitantes				De 2,500 a 49,999 habitantes			
		LI (m.a.s.)	LS (m.a.s.)	LI (Con diseño)	LS (Con diseño)	LI (m.a.s.)	LS (m.a.s.)	LI (Con diseño)	LS (Con diseño)
01	Aguascalientes	0.1094	0.1281	0.0650	0.2140	0.1982	0.2226	0.2044	0.2491
02	Baja California	0.2251	0.2590	0.1681	0.3447	0.3459	0.3759	0.3546	0.4127
03	Baja C. Sur	0.1330	0.1602	0.0923	0.2551	0.3945	0.4249	0.3949	0.4492
04	Campeche	0.0084	0.0141	0.0091	0.0196	0.0905	0.1089	0.0882	0.1278
05	Coahuila	0.0995	0.1107	0.0455	0.0711	0.3120	0.3289	0.3196	0.3577
06	Colima	0.0833	0.1038	0.0608	0.1504	0.2244	0.2563	0.2352	0.3244
07	Chiapas	0.0076	0.0090	0.0073	0.0128	0.0878	0.0949	0.0999	0.1361
08	Chihuahua	0.1157	0.1229	0.0792	0.1310	0.2542	0.2728	0.2612	0.3000
09	Distrito Federal	0.1360	0.1588	0.0949	0.2230	0.2762	0.3028	0.2818	0.3691
10	Durango	0.0438	0.0497	0.0384	0.0654	0.1920	0.2139	0.2182	0.2629
11	Guanajuato	0.0667	0.0727	0.0520	0.1013	0.2616	0.2738	0.2521	0.2953
12	Guerrero	0.0185	0.0213	0.0164	0.0353	0.1662	0.1775	0.1897	0.2207
13	Hidalgo	0.0722	0.0776	0.0470	0.0743	0.2254	0.2375	0.2480	0.2847
14	Jalisco	0.1430	0.1498	0.1462	0.2145	0.3204	0.3293	0.3561	0.3785
15	México	0.0809	0.0855	0.0588	0.0824	0.2187	0.2250	0.2360	0.2566
16	Michoacán	0.0748	0.0796	0.0580	0.0848	0.2170	0.2254	0.2477	0.2745
17	Morelos	0.0995	0.1118	0.0953	0.1564	0.2253	0.2385	0.2616	0.3025
18	Nayarit	0.0411	0.0500	0.0283	0.0722	0.2600	0.2815	0.2858	0.3312
19	Nuevo León	0.1740	0.1853	0.1381	0.1950	0.3437	0.3605	0.3719	0.4157
20	Oaxaca	0.0151	0.0166	0.0141	0.0208	0.1493	0.1575	0.1790	0.2027
21	Puebla	0.0302	0.0329	0.0217	0.0309	0.1504	0.1574	0.1706	0.1945
22	Querétaro	0.0540	0.0633	0.0338	0.0619	0.1909	0.2114	0.1609	0.2918
23	Quintana Roo	0.0681	0.0853	-0.0363	0.1760	0.1594	0.1853	0.1381	0.1978
24	San Luis Potosí	0.0257	0.0292	0.0155	0.0298	0.1697	0.1844	0.2015	0.2505
25	Sinaloa	0.0915	0.1015	0.0935	0.1283	0.2336	0.2514	0.2532	0.2877
26	Sonora	0.1573	0.1668	0.1078	0.1614	0.2899	0.3046	0.3491	0.3784
27	Tabasco	0.0345	0.0400	0.0180	0.0343	0.1727	0.1853	0.1688	0.2042
28	Tamaulipas	0.0724	0.0803	0.0636	0.0971	0.2842	0.3035	0.2961	0.3370
29	Tlaxcala	0.0752	0.0857	0.0438	0.0927	0.1749	0.1862	0.1894	0.2219
30	Veracruz	0.0300	0.0322	0.0213	0.0522	0.1734	0.1794	0.2187	0.2386
31	Yucatán	0.0173	0.0209	0.0139	0.0230	0.0755	0.0831	0.1057	0.1387
32	Zacatecas	0.0483	0.0537	0.0305	0.0603	0.1924	0.2066	0.2120	0.2455
	NACIONAL	0.0577	0.0587	0.0556	0.0638	0.2122	0.2143	0.2437	0.2511

Tabla 2.7 (Continuación) Intervalos al 95% de confianza considerando e ignorando el diseño muestral por tamaño de localidad en cada entidad federativa

Clave del INEGI	Tamaño de localidad <i>t</i> ENTIDAD FEDERATIVA	De 50,000 a 499,999 habitantes				De 500, 000 a más habitantes			
		<i>LI</i> (m.a.s.)	<i>LS</i> (m.a.s.)	<i>LI</i> (Con diseño)	<i>LS</i> (Con diseño)	<i>LI</i> (m.a.s.)	<i>LS</i> (m.a.s.)	<i>LI</i> (Con diseño)	<i>LS</i> (Con diseño)
01	Aguascalientes	---	---	--	--	0.5041	0.5237	0.4966	0.5392
02	Baja California	0.5830	0.6096	0.5956	0.6484	0.6265	0.6386	0.6241	0.6516
03	Baja C. Sur	0.5371	0.5751	0.5487	0.6144	---	---	--	--
04	Campeche	0.3691	0.3933	0.3427	0.3919	---	---	--	--
05	Coahuila	0.4114	0.4329	0.4204	0.4606	0.5178	0.5334	0.5243	0.5585
06	Colima	0.4547	0.4782	0.4475	0.5133	---	---	--	--
07	Chiapas	0.3226	0.3388	0.3171	0.3690	---	---	--	--
08	Chihuahua	0.4729	0.4990	0.4713	0.5135	0.5461	0.5580	0.5259	0.5506
09	Distrito Federal	0.6886	0.6951	0.6791	0.7034	0.675	0.6807	0.6579	0.6751
10	Durango	0.4766	0.4956	0.4732	0.5128	---	---	--	--
11	Guanajuato	0.4233	0.4375	0.4268	0.4722	0.4697	0.4865	0.4685	0.5030
12	Guerrero	0.3552	0.3775	0.3727	0.4145	0.3601	0.3787	0.3678	0.4221
13	Hidalgo	0.4395	0.4643	0.4289	0.5138	---	---	--	--
14	Jalisco	0.4473	0.4608	0.4393	0.4877	0.6589	0.6678	0.6567	0.6780
15	México	0.4676	0.4755	0.4588	0.4857	0.5534	0.5609	0.5410	0.5711
16	Michoacán	0.3816	0.3986	0.3893	0.4358	0.5085	0.5300	0.5055	0.5537
17	Morelos	0.4983	0.5163	0.4751	0.5337	---	---	--	--
18	Nayarit	0.4532	0.4848	0.4477	0.5353	---	---	--	--
19	Nuevo León	0.5862	0.5988	0.5891	0.6173	0.6506	0.6616	0.6414	0.6645
20	Oaxaca	0.3479	0.3696	0.3308	0.4037	---	---	--	--
21	Puebla	0.3211	0.3431	0.3317	0.3854	0.5118	0.5258	0.4973	0.5476
22	Querétaro	0.4077	0.4513	0.3901	0.4836	0.5389	0.5585	0.5413	0.5942
23	Quintana Roo	0.3904	0.4103	0.3329	0.4126	---	---	--	--
24	San Luis Potosí	0.3290	0.3558	0.3306	0.4000	0.5382	0.5585	0.5287	0.5758
25	Sinaloa	0.4952	0.5140	0.5102	0.5559	0.5595	0.5809	0.5732	0.6101
26	Sonora	0.4623	0.4794	0.4786	0.5120	0.5309	0.5527	0.5363	0.5743
27	Tabasco	0.4180	0.4402	0.4053	0.4672	---	---	--	--
28	Tamaulipas	0.4593	0.4702	0.4540	0.4769	---	---	--	--
29	Tlaxcala	0.3684	0.4099	0.3677	0.4756	---	---	--	--
30	Veracruz	0.4240	0.4353	0.4297	0.4651	---	---	--	--
31	Yucatán	---	---	--	--	0.5021	0.5214	0.4904	0.5553
32	Zacatecas	0.4243	0.4502	0.4105	0.4808	---	---	--	--
	NACIONAL	0.4939	0.4969	0.4879	0.4973	0.5941	.05971	0.5876	.5963

Figura 2.1 Intervalos al 95% de confianza por tamaño de localidad ignorando el diseño muestral (m.a.s.)

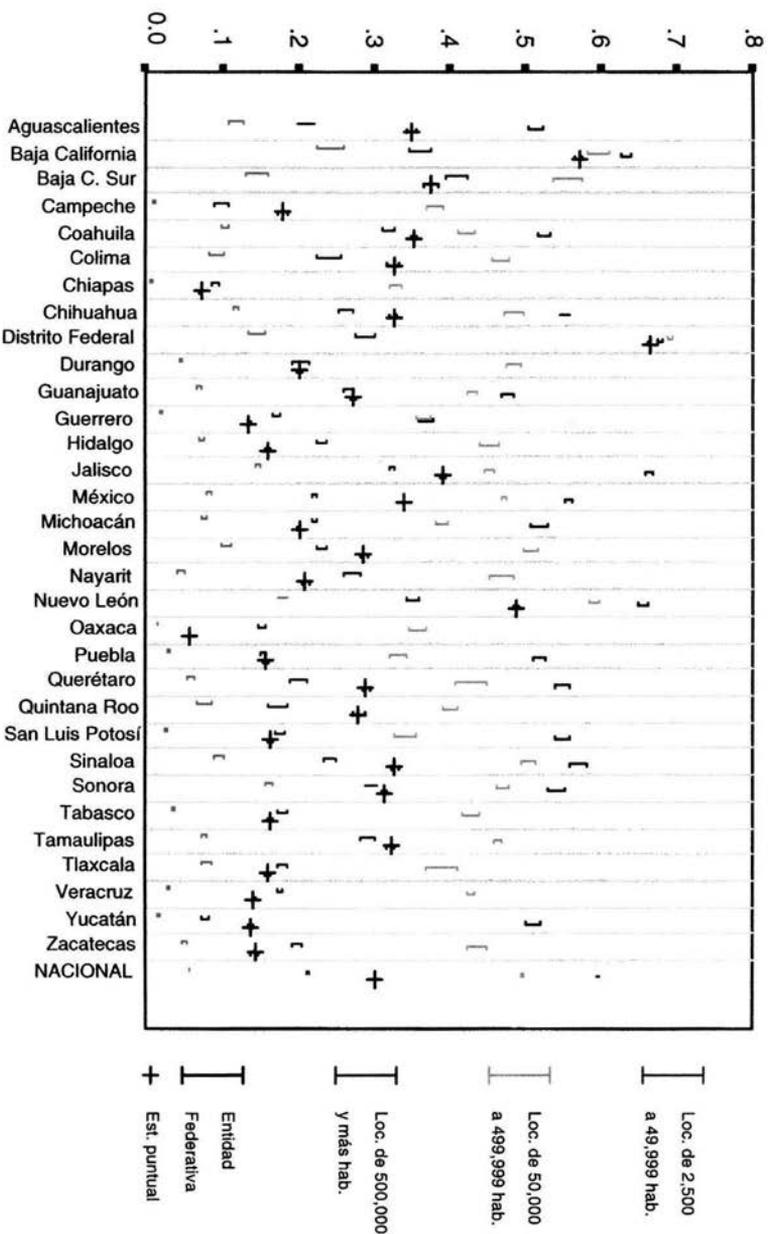
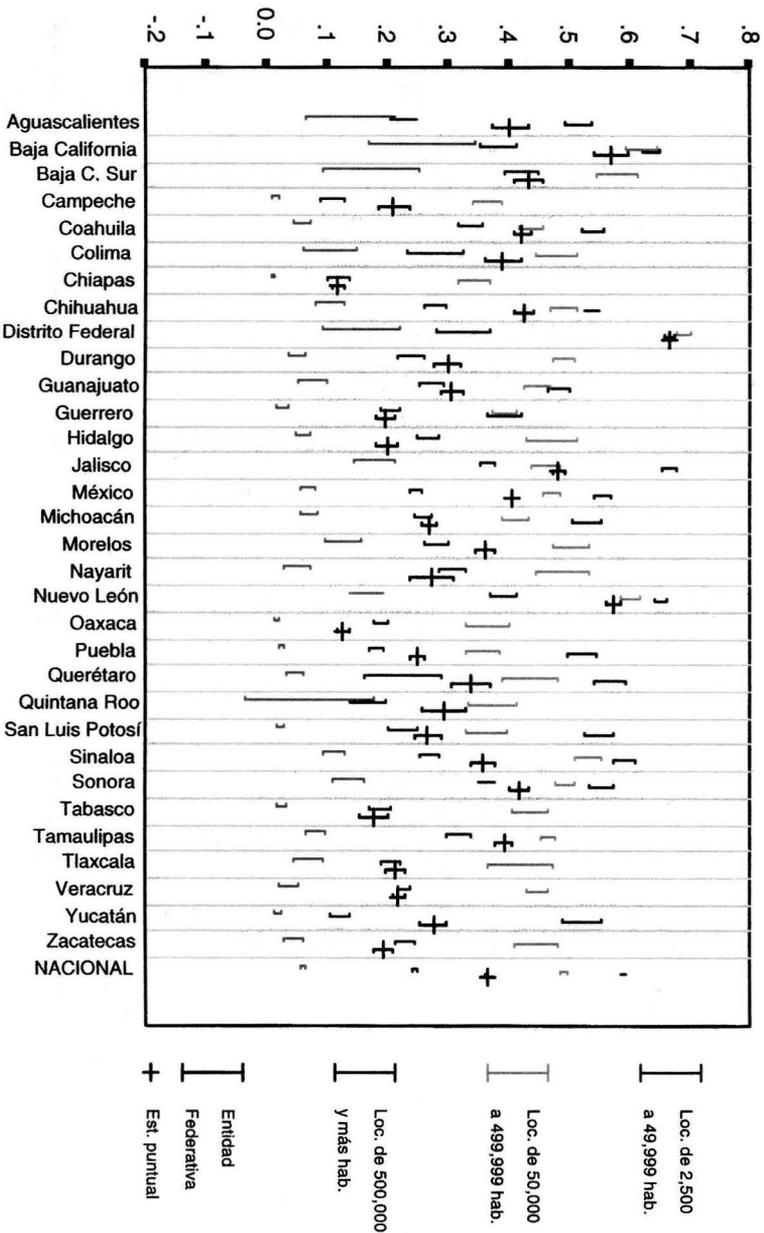


Figura 2.2 Intervalos al 95% de confianza por tamaño de localidad considerando el diseño muestral



Conclusiones

Esta tesis consistió en comparar la magnitud y precisión de los estimadores de la razón de hogares cuya vivienda cuenta con teléfono a partir de la muestra del XII Censo General de Población y Vivienda 2000 considerando e ignorando el diseño de la misma. La proporción poblacional de hogares cuya vivienda particular cuenta con teléfono fue estimada con estimadores de razón, ya que una proporción es un caso particular de una razón.

En la tabla 1.3 del capítulo 1 al observar la comparación de los estimadores de razón, se puede concluir que cuando se considera el diseño muestral en la entidad federativa se obtienen estimaciones más aproximadas al parámetro poblacional que cuando se ignora. A nivel nacional, el estimador de razón ignorando el diseño es de 0.30, mientras que cuando se considera el diseño es de 0.36. En la tabla 1.4 se pueden ver las diferencias porcentuales entre la razón poblacional y el estimador de razón considerando e ignorando el diseño, las diferencias entre la razón poblacional y el estimador de razón considerando el diseño son prácticamente nulas, ya que el promedio de estas diferencias en todos los estados es de aproximadamente 0.5% y a nivel nacional es de 0.14%, mientras que cuando se compara la razón poblacional y el estimador de razón en el que se ignora el diseño el promedio de las diferencias en todos los estados es de aproximadamente 6.5%, y en el caso de Yucatán la diferencia es de 14.7% y a nivel nacional esta diferencia es de 6.3%. Con lo anterior se muestra que para

obtener la mejor estimación del verdadero parámetro, es importante considerar el diseño muestral, esto es sabido, pero aquí se cuantifican las diferencias.

Al ignorar el diseño muestral, las proporciones estimadas se alejan de las proporciones poblacionales. Al considerar el diseño muestral, las proporciones estimadas se acercan más a las proporciones poblacionales, pero su varianza es mayor que la obtenida al ignorar el diseño.

En la Tabla 1.5 del capítulo 1 se puede observar que el efecto de diseño (Deff) en las entidades federativas es considerable, ya que como se había mencionado antes, éste varía de 6.38 para Baja California Sur a 62.31 para el caso de Oaxaca, y a nivel nacional es de 25.71. El Deff nos indica el número de veces en que la varianza bajo el diseño de la muestra es mayor o menor que la correspondiente a la varianza bajo m.a.s. En este trabajo se observó que la primer varianza es mayor que la segunda, ya que cuando conglomeramos se pierde precisión en ésta. Continuando con la Tabla 1.5 se añadió una columna en ésta, que es el cociente del error estándar del estimador de razón considerando diseño entre el error estándar del estimador de razón bajo m.a.s., ya que en el anexo del Diseño de la muestra el INEGI así define al Deff.

Por lo anterior, es importante que cuando se realice la estimación de algún parámetro poblacional a través de una muestra, utilicemos el tipo de muestreo con el cual la estimación puntual sea lo más cercano al parámetro y esto último se logró considerando el diseño muestral.

Al final del capítulo 1 se puede observar que los intervalos al 95% de confianza ignorando el diseño muestral son más pequeños que cuando se ignora el diseño.

Lo que se podría concluir de los resultados del segundo capítulo, es muy similar a lo que se concluyó del primer capítulo, la diferencia es que aquí, en lugar de trabajar con las entidades federativas completas, se trabajó con tamaños de localidad. En el capítulo 2 la diferencia porcentual entre el estimador de razón considerando el diseño muestral y en el que se ignora el diseño no es muy significativa, ya que la diferencia que existe por entidad federativa completa se divide en 3 ó 4 tamaños de localidad, pero las estimaciones de razón siguen siendo más aproximadas al verdadero parámetro cuando se considera el diseño muestral que cuando se ignora.

Por último, al ver la Tabla 2.7 y las Figuras 2.1 y 2.2 se concluye que entre mayor es el grado de urbanización de los tamaños de localidad, mayores son los valores que toman los estimadores, es decir, que las localidades más urbanizadas cuentan con una mayor proporción de hogares que cuentan con teléfono.

Bibliografía

COCHRAN, W.G. (1977), Sampling Techniques, third edition, John Wiley & Sons, E.U.A.

Contar 2000, Sistema para la consulta de tabulados y bases de datos de la muestra, INEGI.

KISH, L. (1979), Muestreo de Encuestas, segunda edición en español, Trillas.

PC CARP, IV (Program Algorithms), Statistical Laboratory, Iowa State University, Ames, 1986, , págs. 74-76.

Sitios de Internet visitados

Se visitaron varias páginas web cuya liga original fue la siguiente:

<http://www.inegi.gob.mx> Instituto Nacional de Estadística, Geografía e Informática.

Anexo 1

Diseño de la muestra

Este Anexo del Diseño de la muestra es una versión aumentada, obtenida a partir de los Anexos de Contar 2000, Sistema para la consulta de tabulados y base de datos de la muestra del XII Censo General de Población y Vivienda 2000, INEGI.

Las unidades de análisis de la muestra del XII Censo General de Población y Vivienda 2000 (XII CGPV-2000) fueron seleccionadas con criterios probabilísticos (se mencionaran dentro del diseño) con el objeto de realizar estimaciones de los indicadores de interés de la población bajo estudio y al mismo tiempo conocer la calidad de las mismas. Los indicadores a estimar son fundamentalmente tasas, promedios, razones y proporciones. El procedimiento de selección de la muestra permite generar información con desagregación en la mayor parte de los indicadores.

Para el levantamiento del censo y de la muestra del XII CGPV-2000 se utilizaron el cuestionario básico y el ampliado respectivamente. El conjunto de viviendas de la muestra es un subconjunto de las viviendas para ser censadas. A cada vivienda se le visitó sólo una vez. El cuestionario ampliado se generó de tal manera que incluyera, en su totalidad, las preguntas del cuestionario básico y algunas mas. De tal forma que a las viviendas en muestra que se les aplicó el cuestionario ampliado también se les consideró censadas. Para la generación de las bases de explotación del censo en el caso de las viviendas en muestra, se consideraron sólo las respuestas a las preguntas comunes al cuestionario básico.

Esquema de Muestreo

El esquema de muestreo es por conglomerados y en una sola etapa, es decir, se seleccionan áreas geográficas completas bien determinadas, estas últimas también conocidas como unidades primarias de muestreo pueden estar conformadas por una manzana, un grupo de manzanas, una localidad, un grupo de localidades, un ageb, un grupo de ageb, y hasta un municipio según el esquema aplicado a cada región.

En la muestra entraron con certeza todos los municipios del país, así como aquellas localidades que de acuerdo con la cifra obtenida durante el Censo de Población y Vivienda 1995 contaban con una población mayor o igual a 2000 habitantes.

Marco muestral

El marco de muestreo está constituido por los catálogos¹ de Ageb, manzanas y localidades, así como la información estadística asociada a cada área y el material cartográfico que permite su identificación y ubicación en campo.

Para la definición del marco muestral se tomó como base la información generada por el Censo de Población y Vivienda 1995 con actualizaciones en los años 1998 y 1999 previo al censo, y en el 2000, durante el censo para la parte rural, y por la Enumeración Integral 1998 tanto para la parte urbana como para las localidades rurales ameznadas.

Las unidades de muestreo varían de acuerdo con el tipo de área y el diseño muestral que se aplica en cada una de ellas.

¹ Catálogo se refiere a una lista de claves de Ageb.

- **Área urbana y localidades ameznadas de 2,000 a 2, 499 habitantes.** La Unidad Primaria de Muestreo (UPM) está constituida generalmente por manzanas, aunque en algunas localidades con baja densidad de población, la UPM está conformada por un Ageb urbana completa.
- **Área rural.** Las UPM pueden ser localidades o Ageb rurales.

Tamaño de muestra y precisión

Para garantizar que las estimaciones que se obtengan a partir de la muestra del XII CGPV-2000 tengan calidad aceptable, fue necesario, entre otros elementos, que el tamaño de muestra que se definió para cada parámetro de interés fuera suficiente.

Para la determinación de dicho tamaño se elige habitualmente el método de afijarla con un indicador de interés pero de poca frecuencia, es decir, raro en la población. Esto es para tener la seguridad de que si se está fijando un tamaño de muestra que permita estimar con cierta calidad y precisión un parámetro como éste; el mismo tamaño de muestra será suficiente para garantizar al menos ese nivel de calidad y precisión para la mayoría de los indicadores a estimar. En este caso se seleccionó el indicador "hogares con preceptores de ingresos por otras fuentes distintas al salario", incluido dentro de las variables del cuestionario ampliado.

El tamaño de muestra se obtuvo mediante la siguiente expresión:

$$n_0 = \frac{Q}{P} * \frac{z^2}{d^2} * \frac{DEFT}{1 - TNR} \quad (1)$$

donde:

P = Proporción a estimar.

$$Q = 1 - P$$

z = Valor asentado en tablas de la normal estándar que garantiza realizar estimaciones con determinada confianza.

d = Error relativo máximo aceptable, éste indica que con el tamaño de muestra calculado, se espera que los errores relativos de la mayoría de los indicadores a estimar no pasen de dicha magnitud.

DEFT = Efecto de diseño, que es la pérdida o ganancia en la eficiencia del diseño por el efecto de conglomerar elementos de la población para formar unidades muestrales.

TNR = Tasa de no respuesta. Este valor esperado se expresa respecto al total planeado de viviendas en la muestra. Es una especulación basada en la experiencia de eventos anteriores para asegurar que los tamaños de muestra esperados no resulten insuficientes por causa de la no respuesta voluntaria o involuntaria de los informantes.

Los valores que tomaron los indicadores en (1) son:

P = 0.2386 Proporción de hogares con preceptores de ingresos por otras fuentes distintas al salario (valor obtenido del Censo de Población y Vivienda 1995).

$$Q = 1 - 0.2386 = 0.7614.$$

z = 1.645 valor del cuantil en tablas de la normal estándar correspondiente a una confianza del 90%.

d = 0.1427 Error relativo máximo aceptable.

DEFT = 1.44 Efecto de diseño esperado (valor derivado de la encuesta del Censo de Población y Vivienda 1995).

TNR = 0.15 Tasa de no respuesta máxima de 15%.

De esta manera se obtuvo un tamaño de muestra de 718 hogares. Al dividir esta cantidad entre el promedio de hogares por vivienda, se obtuvo un tamaño de muestra en términos de viviendas igual a 700. Como ya se había mencionado, lo anterior garantiza obtener estimaciones a nivel municipal con calidad aceptable, lo mismo que en localidades de 50,000 y más habitantes.

Debido a que varios municipios cuentan con un número de viviendas aproximado e incluso menor al tamaño de muestra descrito, se realizó un ajuste por población finita al tamaño de muestra mínimo requerido.

Así, para estados como Chiapas, Hidalgo, Jalisco y Veracruz, el tamaño de muestra mínimo por municipio se fijó en 500 viviendas, mientras que en Oaxaca, Yucatán, Puebla y Tlaxcala la muestra mínima por municipio varía entre 150 y 250 viviendas. En estas entidades, los municipios con menos de 1,000 habitantes entraron completos a la muestra, es decir, fueron censados a través del cuestionario ampliado.

Uno de los motivos por los que se tomaron completos algunos municipios, es por que en ciertos casos, resultaba más costoso en tiempo, dinero, planeación y operación levantar una muestra del municipio que levantar el censo del mismo.

El ajuste por población finita se calculó de la siguiente manera:

$$n = \frac{n_0}{1 + \frac{n_0}{N}} \quad (2)$$

donde:

n = Tamaño de muestra ajustado por población finita.

n_0 = Tamaño de muestra original.

N = Tamaño de viviendas por municipio.

Aplicando el promedio de habitantes por viviendas (para el estado de Oaxaca el promedio fue de 4.9 de acuerdo con los resultados del Censo de Población y vivienda 1995) a municipios con 1,000 habitantes se tiene:

$$N = \frac{1,000 \text{ habitantes}}{4.9 \text{ habitantes/viviendas}} = 204 \text{ viviendas}$$

Sustituyendo este resultado en (2) se tiene:

$$n = \frac{700}{1 + \frac{700}{204}} = 158$$

que para casos prácticos se redondeó a 150.

Procedimiento de selección

El procedimiento de selección de la muestra del XII CGPV-2000 es distinto para cada municipio, dependiendo del tipo de área (urbana o rural).

Área urbana: Aunque la mayoría de las Ageb tienen muestra, sólo en algunas localidades urbanas con menos de 5,000 habitantes y baja densidad de población, esta regla no se cumple.

Debido a esto, se realizaron dos tipos de selección en el interior de las localidades urbanas:

- **Selección de manzanas en el interior de cada Ageb**

Se llevó a cabo una selección de por lo menos dos manzanas en cada Ageb. la selección se realizó a través de un m.a.s.

Las Ageb con dos o menos manzanas, entraron completas a la muestra: En los casos en donde los municipios o localidades no alcanzaron el mínimo requerido, se seleccionaron más manzanas por Ageb.

- **Selección de Ageb completas en el interior de la localidad**

En algunas localidades urbanas menores a 5,000 habitantes y con un promedio de viviendas por manzana inferior a cinco, se optó por la selección de áreas cuya identificación en campo fuera más fácil (Ageb), ya que en ocasiones la traza irregular o poco definida de las manzanas impidió identificar las áreas seleccionadas.

La selección de Ageb se realizó mediante un m.a.s.

Área rural. Las localidades de 2,000 a 2,499 habitantes entraron con certeza a la muestra. Si éstas contaban con manzanas regulares, entonces se les trataba como localidad urbana, es decir, se seleccionaron dos o más manzanas en su interior; en cambio, cuando no contaban con plano, entraban completas a la muestra.

En el resto del área rural (localidades menores a 2,000 habitantes) la selección dependió de las características de cada municipio, por lo que se aplicaron tres esquemas de selección.

- **Selección de localidades por cada Ageb del municipio**

Al igual que en el área urbana, esta selección se caracteriza por una gran dispersión de la muestra, lo cual se traduce en una mejor calidad en la precisión de los estimadores.

Se escogieron dos o más localidades de cada Ageb (dependiendo del mínimo de viviendas requerido para el municipio). La selección fue a través de un m.a.s.

- **Selección de localidades en el interior del municipio**

Al tratarse de un muestreo por áreas, se pierde el control sobre el tamaño de muestra obtenido, ya que éste depende del número de viviendas de las áreas seleccionadas.

Con la finalidad de controlar y reducir el tamaño cuando éste sobrepasa con mucho al requerido, se aplicó otra selección que consistió en estratificar las localidades rurales en el interior del municipio según el total de habitantes; en el interior de cada estrato conformado se hizo una selección de algunas localidades, de acuerdo con la Tabla A1.1.

Estrato	Rangos de tamaños de localidad (habitantes)	Número de localidades a seleccionar
1	Con menos de 50	4
2	De 50 a 499	2
3	De 500 a 1999	1

Tabla A1.1 Localidades a seleccionar de acuerdo al estrato.

La selección en el interior de cada estrato se realizó con igual probabilidad y mediante muestreo sistemático.

- **Selección de Ageb (rurales) en el interior del municipio**

En los casos en los que los municipios cuentan con un gran número de localidades rurales, y el promedio de viviendas por localidad fue menor a 15, se seleccionaron Ageb rurales completas. Dicha selección se realizó con probabilidad igual y mediante m.a.s.

Para los tres esquemas anteriores, las Ageb que tienen menos de X viviendas ($50 \leq X \leq 150$ dependiendo de la dispersión de la entidad) entran completas a la muestra. Esto último, como ya se había mencionado anteriormente se realizó entre otros motivos cuando el levantamiento de la muestra resulta ser más costosa, en dinero, planeación, operación y tiempo.

Procedimiento de selección

De acuerdo con el proceso de selección, se tienen dos métodos para el cálculo de la probabilidad de que las viviendas sean seleccionadas.

El primer caso (en el que se realiza un muestreo aleatorio simple) considera la selección de manzanas en el interior de cada Ageb urbana, selección de Ageb completas en el interior de la localidad urbana, selección de localidades por cada Ageb rural del municipio y selección de Ageb rurales en el interior del municipio. La fórmula que se empleó es la siguiente:

$$P_{ij} = \frac{n_i}{N_i}$$

donde:

P_{ij} = Probabilidad de selección de la j -ésima UPM en el ámbito de selección correspondiente i .

n_i = Número de UPM (manzanas, localidades o Ageb) a seleccionar.

N_i = Total de UPM en el ámbito de selección (Ageb urbana o rural, municipio).

n_i y N_i varían según el tipo de selección que se realice, de acuerdo a la Tabla A1.2 .

Tipo de selección	n_i	N_i
Área urbana		
Selección de manzanas en el interior de cada Ageb.	Número de manzanas a seleccionar en la i -ésima Ageb.	Total de manzanas en la i -ésima Ageb.
Selección de Ageb completas en el interior de la localidad urbana.	Número de Ageb a seleccionar en la i -ésima localidad urbana.	Total de Ageb en la i -ésima localidad urbana.
Área rural		
Selección de manzanas por localidad amezanada de 2,000 a 2,499 habitantes.	Número de manzanas a seleccionar en la i -ésima localidad rural.	Total de manzanas en la i -ésima localidad rural.
Selección de localidades por cada Ageb del municipio.	Número de localidades a seleccionar en la i -ésima Ageb.	Total de localidades en la i -ésima Ageb.
Selección de Ageb en el interior del municipio.	Número de Ageb a seleccionar en el i -ésimo municipio.	Total de Ageb en el i -ésimo municipio.

Tabla A1.2. Unidades para la probabilidad de selección.

En el segundo caso (en el que se utilizó muestreo sistemático) es cuando en el área rural se utiliza una selección de localidades en el interior del municipio, y la expresión que se utiliza es:

$$P_{hi} = \frac{l_h}{L_h}$$

donde:

P_{hi} = Probabilidad de selección de la i -ésima localidad del estrato h en el municipio; todas las viviendas que conforman las localidades seleccionadas tienen la misma probabilidad de selección.

l_h = Número de localidades a seleccionar en el estrato h del municipio.

L_h = Total de localidades en el estrato h del municipio.

Ámbitos de selección

El ámbito de selección está sujeto a diversos criterios dependiendo de la situación geográfica y poblacional en cada entidad, municipio, zona o localidad del país.

Los ámbitos de selección se forman de acuerdo al tamaño de localidad en el municipio; en una parte de la entidad o en un grupo de municipios, según los tamaños de los municipios en la entidad (entre otros criterios) y pueden ser los siguientes:

- 1.- Localidades con 500,000 y más habitantes.
- 2.- Grupo de localidades con más de 100,000 habitantes.
- 3.- Grupo de localidades de 50,000 a 99,999 habitantes.
- 4.- Grupo de localidades de 20,000 a 49,999 habitantes.
- 5.- Grupo de localidades 15,000 a 19,999 habitantes.
- 6.- Grupo de localidades 2,500 a 14,999 habitantes.
- 7.- Grupo de localidades 2,000 a 2,499 habitantes.

- 8 .- Grupo de localidades 500 a 1,999 habitantes.
- 9 .- Grupo de localidades 50 a 499 habitantes.
- 10 .- Grupo de localidades con menos de 50 habitantes.
- 11 .- Un Ageb.
- 12 .- La parte rural de un municipio.
- 13 .- Un municipio.

Estimadores

La construcción de estimadores para los totales de la población se hizo en dos etapas:

Primero se obtuvieron estimaciones de totales a partir de los ponderadores (inverso de la probabilidad de selección).

En general, el estimador del total municipal obtenido directamente de la muestra es:

$$\hat{Y}_m = \sum_i \left(F_i \cdot \sum_j \sum_k y_{ijk} \right)$$

donde:

F_i = Es el inverso de la probabilidad de selección de las UPM (factor de expansión) correspondientes al ámbito de selección i .

y_{ijk} = Es el valor de la característica en estudio de la k -ésima entrevista de la j -ésima UPM, correspondiente al ámbito de selección i .

Después se utilizó un estimador separado de razón, el cual empleó como variable auxiliar a la población de viviendas particulares habitadas, obtenida de la integración de las cifras preeliminares del XII CGPV-2000.

El estimador separado de razón es:

$$\hat{Y}_{Rm} = \frac{\hat{Y}_m}{\hat{X}_m} X_m$$

donde:

\hat{Y}_{Rm} = Estimador separado de razón del total de la característica Y obtenido directamente de la muestra para el municipio m .

\hat{Y}_m = Estimador del total de la característica Y obtenido directamente de la muestra para el municipio m .

\hat{X}_m = Estimador del total de la población obtenido directamente de la muestra para el municipio m .

X_m = Número de residentes generado a partir de cifras preeliminares.

Para la estimación de promedios, tasas y proporciones se utilizaron estimadores de razón combinados, definiendo otra característica T de manera similar a Y , quedando entonces:

$$\hat{R} = \frac{\hat{Y}}{\hat{T}}$$

Precisión y confianza

Debido a que los datos presentados se estimaron a partir de la muestra, éstos pueden diferir de los resultados obtenidos a partir del Censo.

Dicha diferencia entre el indicador muestral y el censal puede cuantificarse y estimarse a partir de la confianza y precisión de los estimadores, para los cuales se utilizaron las siguientes expresiones.

Error estándar (EE). Para el cálculo del error estándar de estimadores de razón del diseño de la muestra se utilizó la siguiente fórmula:

$$EE = \sqrt{\text{Var}(\hat{R})} = \sqrt{\frac{1}{\hat{T}^2} \sum_m \sum_i \frac{n_{mi}}{n_{mi}-1} (1-f_{mi}) \left[\sum_j^{n_{mi}} \left(d_{mij} - \frac{d_{mi}}{n_{mi}} \right)^2 \right]}$$

donde:

M = Total de municipios.

I_m = Ámbitos de selección en el municipio m .

n_{mi} = Total de UPM seleccionadas en el ámbito de selección i del municipio m .

f_{mi} = Fracción de muestreo en cada ámbito de selección i de cada municipio m .

$$d_{mij} = \hat{y}_{mij} - \hat{R} \hat{t}_{mij}$$

donde:

\hat{y}_{mij} y \hat{t}_{mij} son los estimadores del total de población de la característica Y y la característica T de la j -ésima UPM en el ámbito de selección i , municipio m , las cuales se calculan mediante la siguiente expresión:

$$\hat{y}_{mij} = \sum_k^{P_{mij}} (F_i y_{mijk})$$

para la j -ésima UPM, ámbito i , municipio m , donde P_{mij} es el total de las unidades de análisis en la UPM y F_i es el factor de expansión de las UPM en el ámbito i .

\hat{t}_{mij} se obtiene de manera análoga a \hat{y}_{mij} .

$d_{mi} = \hat{y}_{mi} - \hat{R} \hat{t}_{mi}$ es la suma para el i -ésimo ámbito de selección en el municipio m , donde:

\hat{y}_{mi} y \hat{t}_{mi} son las estimaciones para los totales de la población con la característica y y t del ámbito de selección i del municipio m .

Coefficiente de variación (CV). Es la medida relativa, definida como el error estándar dividido entre el estimador:

$$CV = \frac{EE}{\hat{R}}$$

Efecto de diseño (DEFT). Es la razón de comparación entre el error estándar de acuerdo con el diseño de la muestra y el error estándar bajo un diseño de m.a.s. ($EE_{m.a.s.}$).

$$DEFT = \frac{EE}{EE_{m.a.s.}}$$

Intervalo de confianza al 90%. El intervalo de confianza indica que el valor real se encuentra entre el límite inferior y el límite superior con confianza del 90%, y se calcula como sigue:

$$\left(\hat{R} - 1.645 \cdot EE, \hat{R} + 1.645 \cdot EE \right).$$

Anexo 2

Distribución del Factor de expansión por entidad federativa y por tamaño de localidad

En este anexo se muestran las gráficas de la distribución del factor de expansión correspondientes a las 32 entidades federativas y a nivel nacional. El comportamiento que se puede observar de éstas es que para los estados en donde la mayoría de las localidades son rurales como lo es el caso de Chiapas la distribución del factor está sesgada hacia la izquierda, resultando una gran frecuencia del factor de expansión que representa a un hogar. Lo anterior se debe a la manera en como se tomó la muestra, ya que en algunas localidades rurales resultó más barato en muchos aspectos levantar un censo que tomar una muestra. En cambio en los estados donde predominan las localidades urbanas el factor de expansión tiende a parecerse al de una distribución normal como lo es el caso del Distrito Federal. Por otra parte, en las gráficas donde se muestra la distribución del factor de expansión por tamaño de localidad (últimas 15 gráficas del Anexo 2 correspondientes a los estados de Baja California, Distrito federal, Sonora y Yucatán) se puede observar que en general el factor de expansión que representa de una a cuatro viviendas, en su mayoría caen dentro de tamaños de localidad $t=1,2$, mientras que los factores de expansión que representan a más de 5 viviendas están distribuidos en los cuatro tamaños de localidad, y la frecuencia de ellos depende del grado de urbanización y del tamaño de muestra que se hizo de cada uno de los tamaños de localidad en cada entidad federativa.

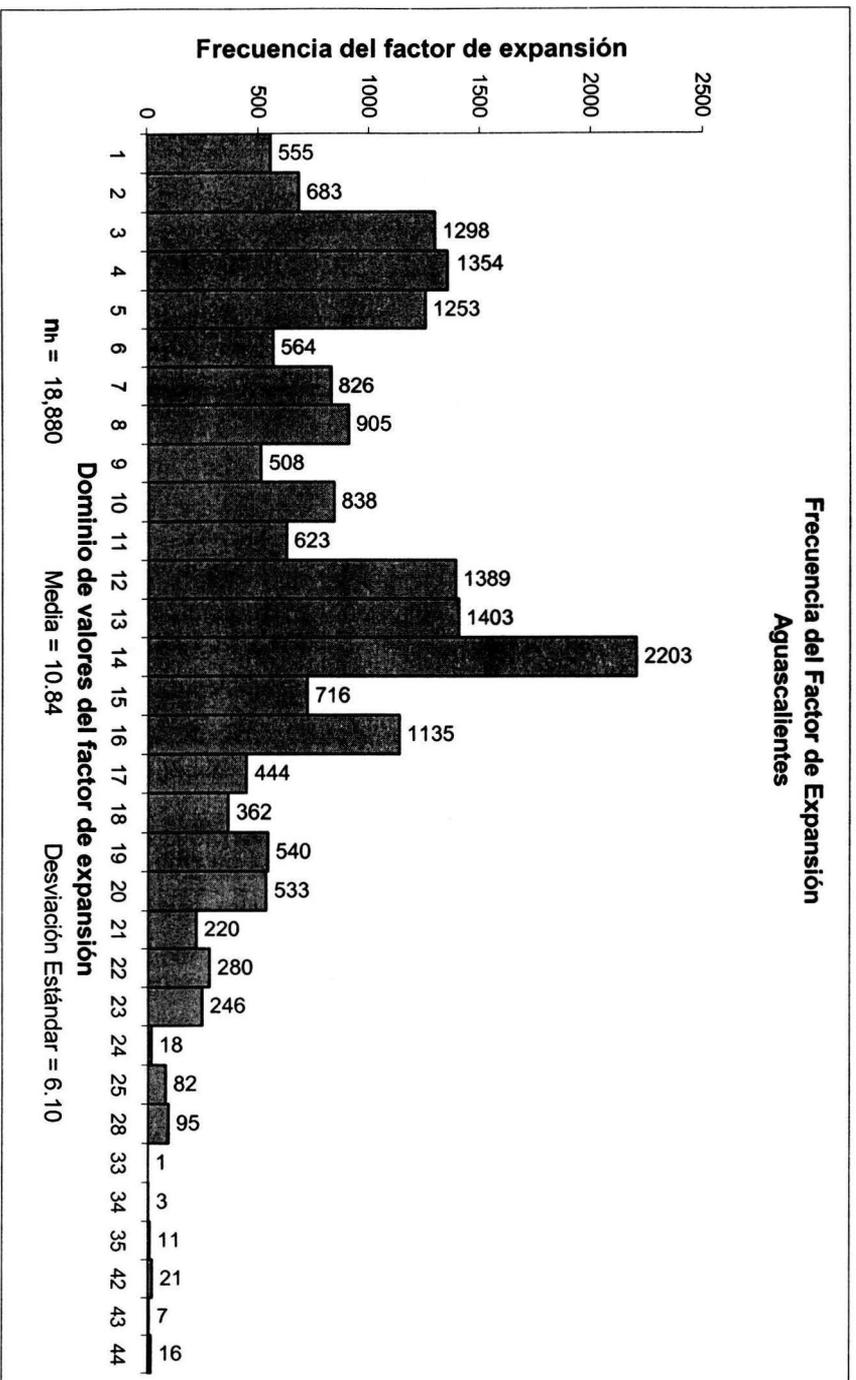


Figura A2.1 Frecuencia de los valores tomados por el Factor de expansión en Aguascalientes

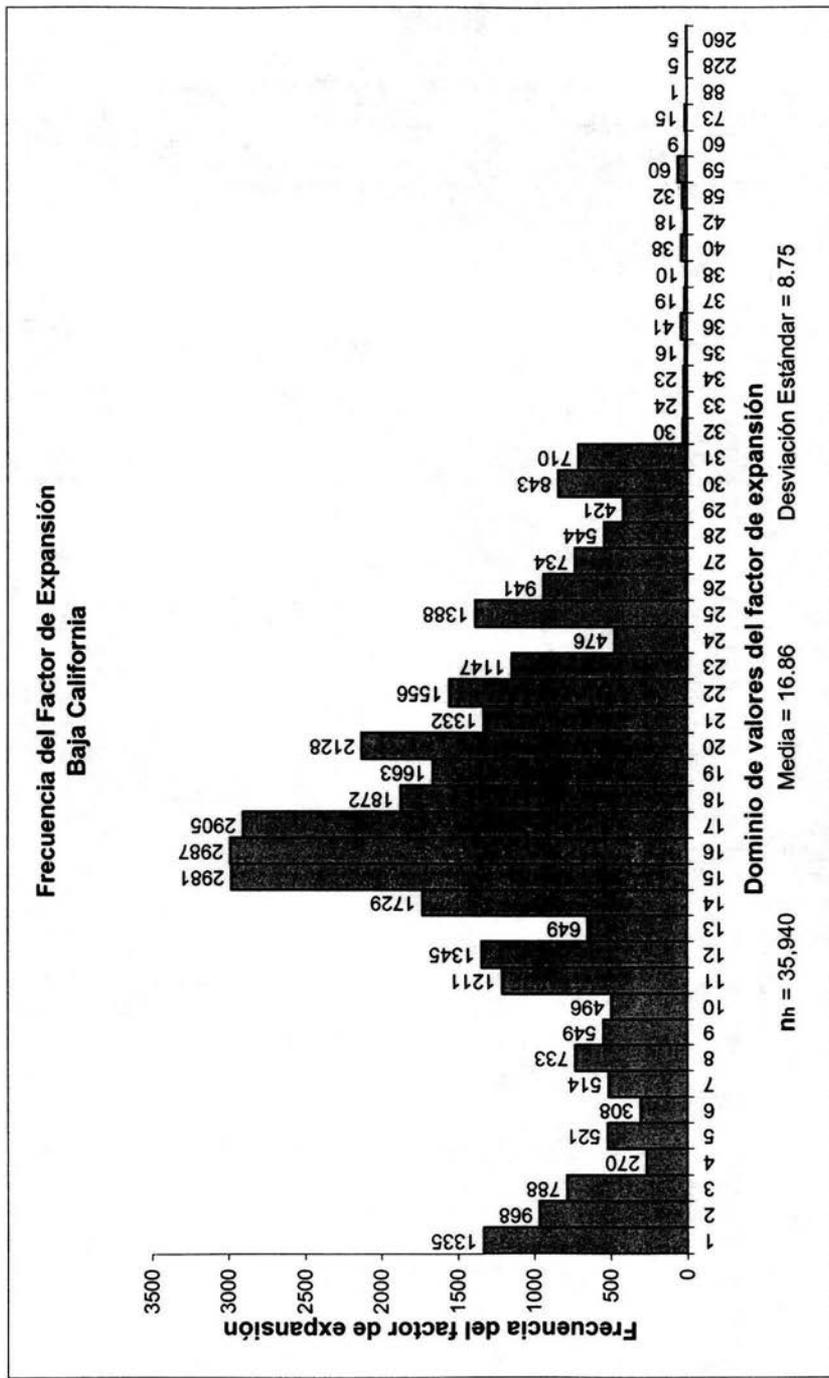


Figura A2.2 Frecuencia de los valores tomados por el Factor de expansión en Baja California

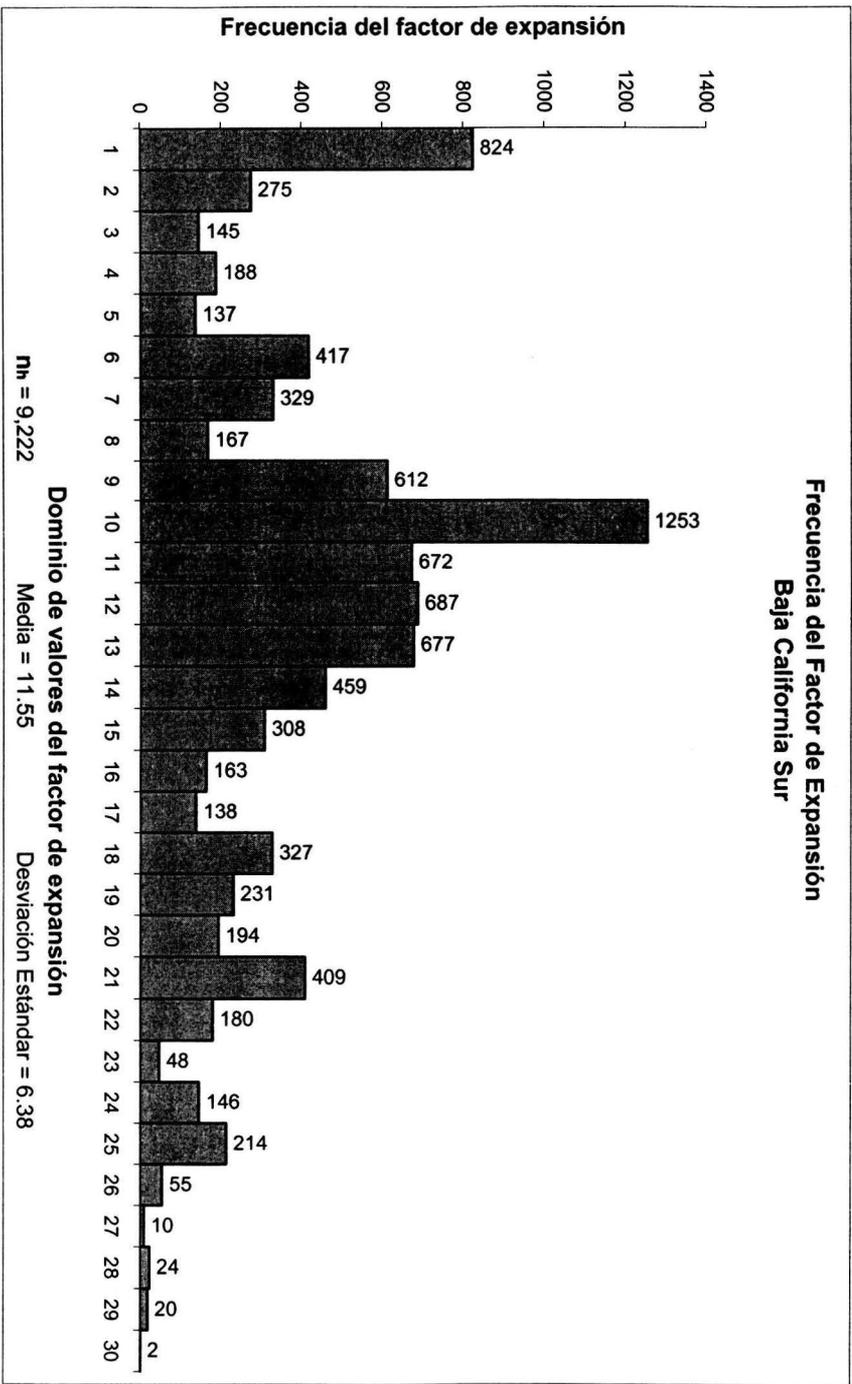


Figura A2.3 Frecuencia de los valores tomados por el Factor de expansión en Baja California Sur

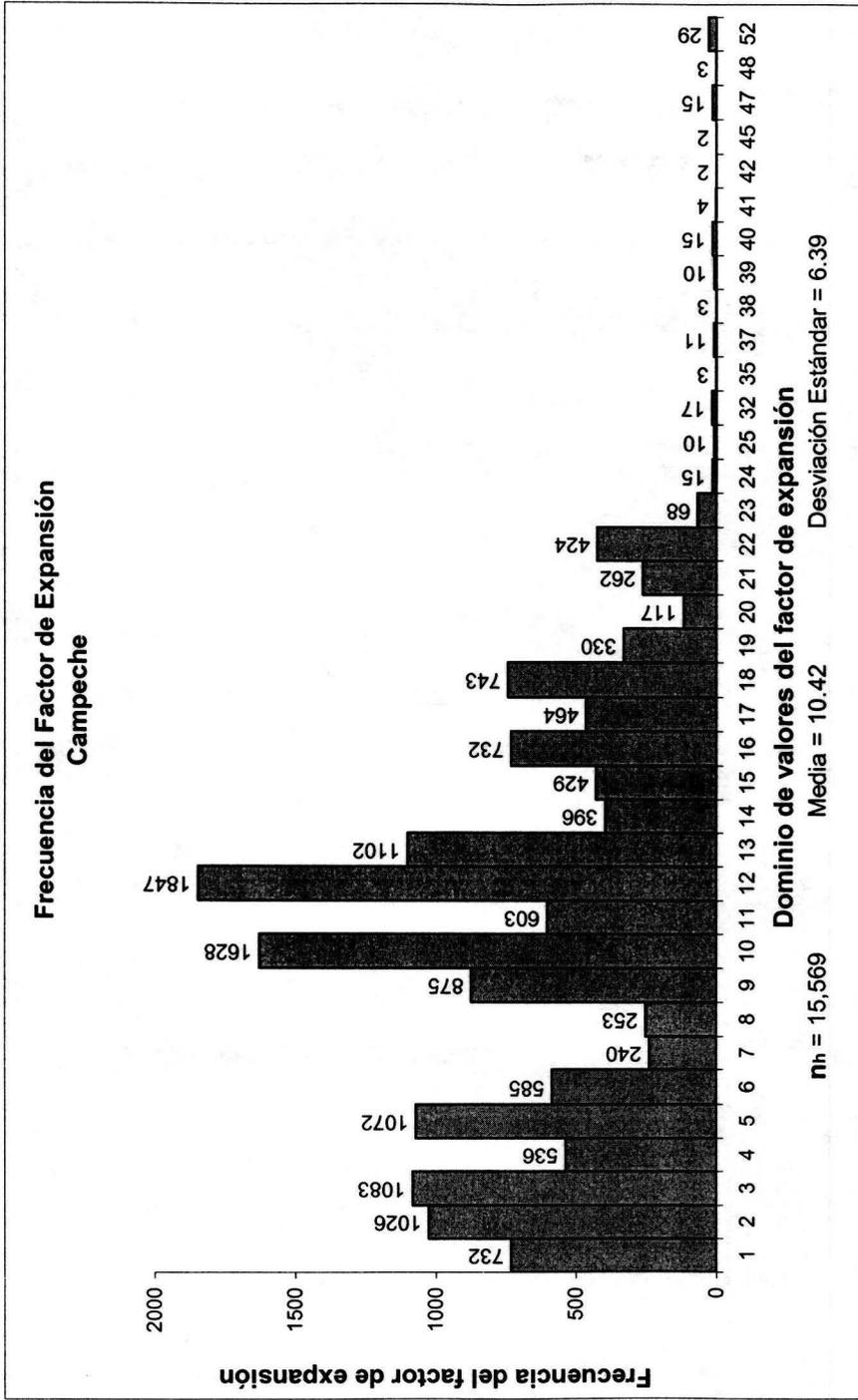


Figura A2.4 Frecuencia de los valores tomados por el Factor de expansión en Campeche

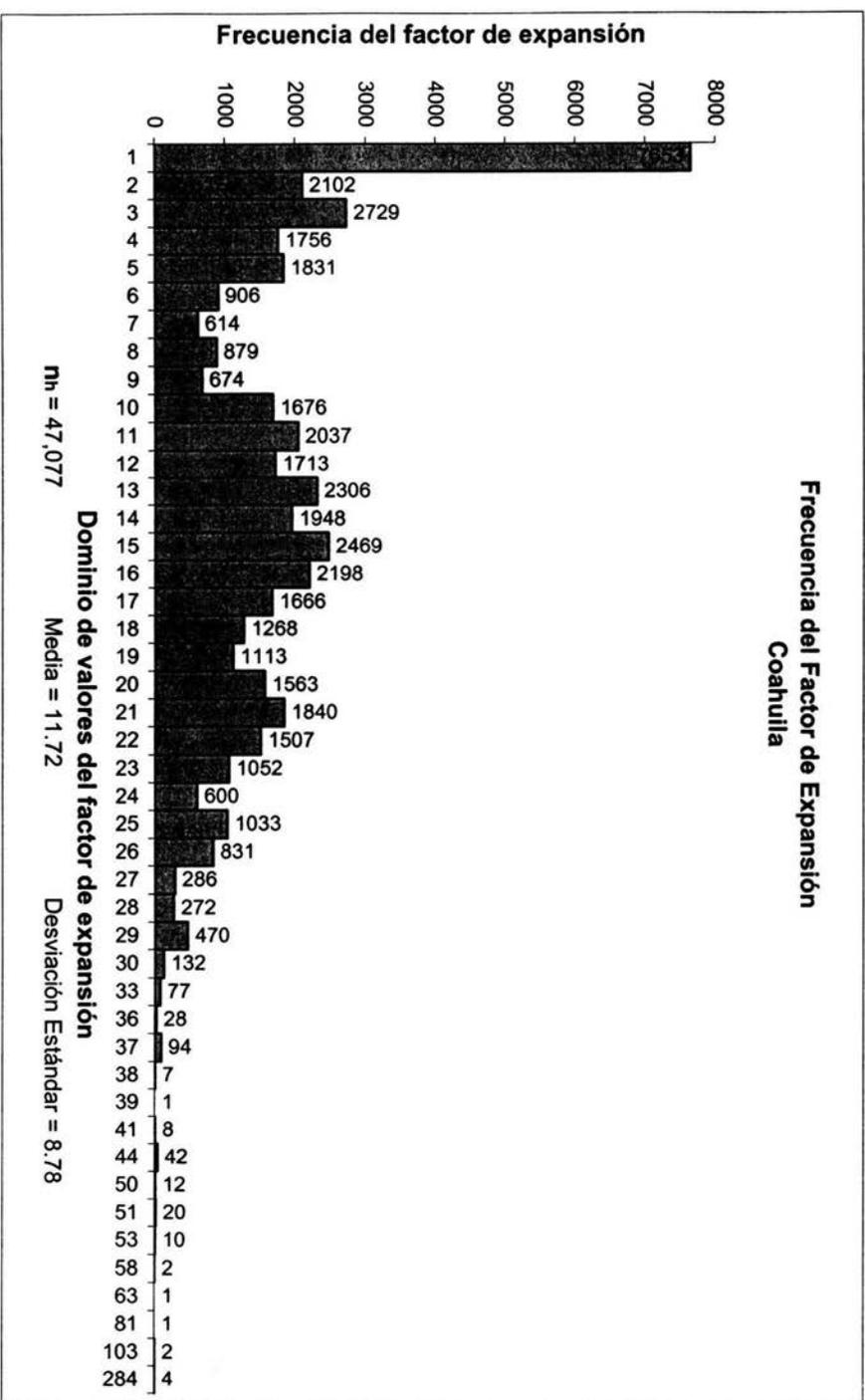


Figura A2.5 Frecuencia de los valores tomados por el Factor de expansión en Coahuila

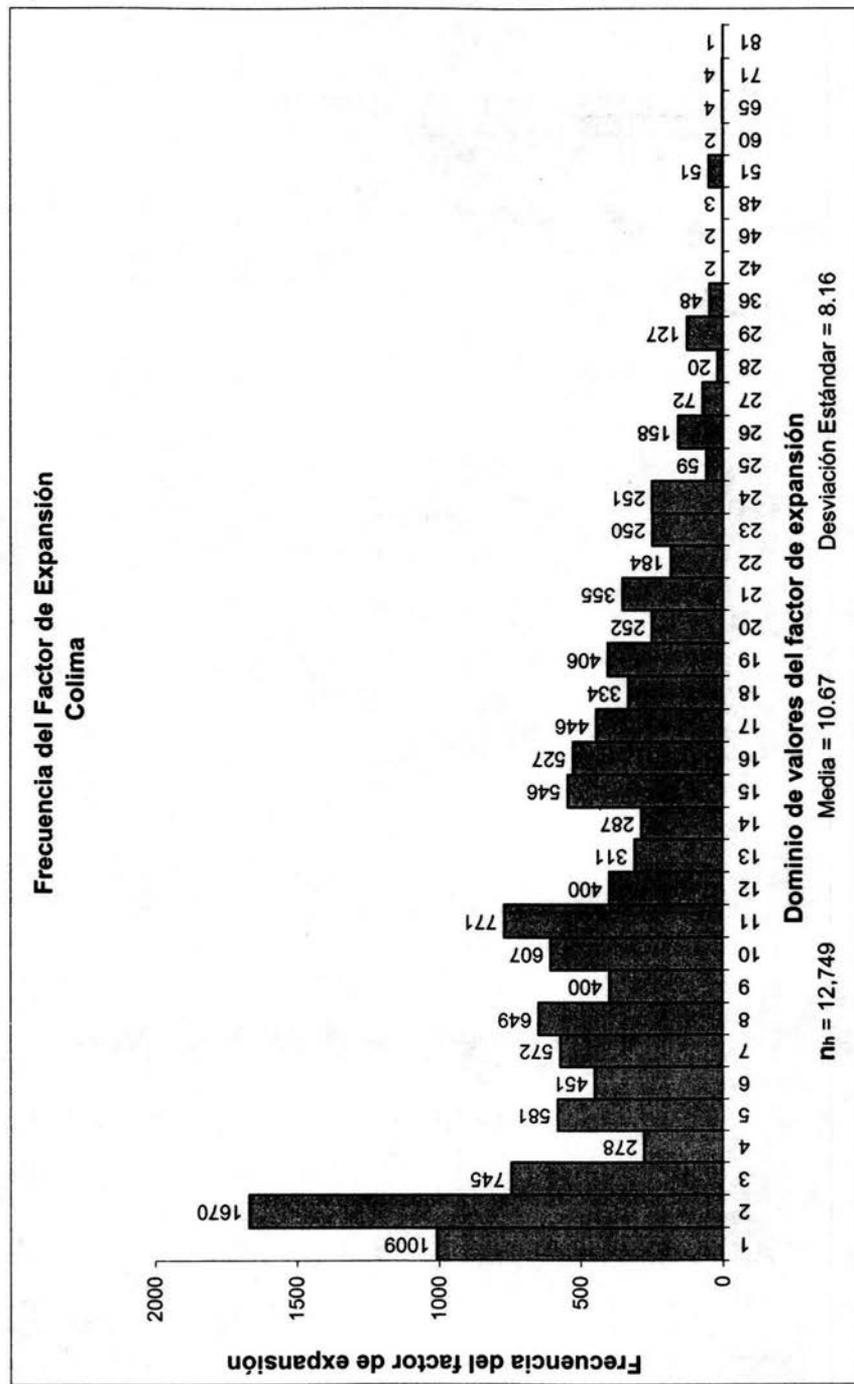


Figura A2.6 Frecuencia de los valores tomados por el Factor de expansión en Colima

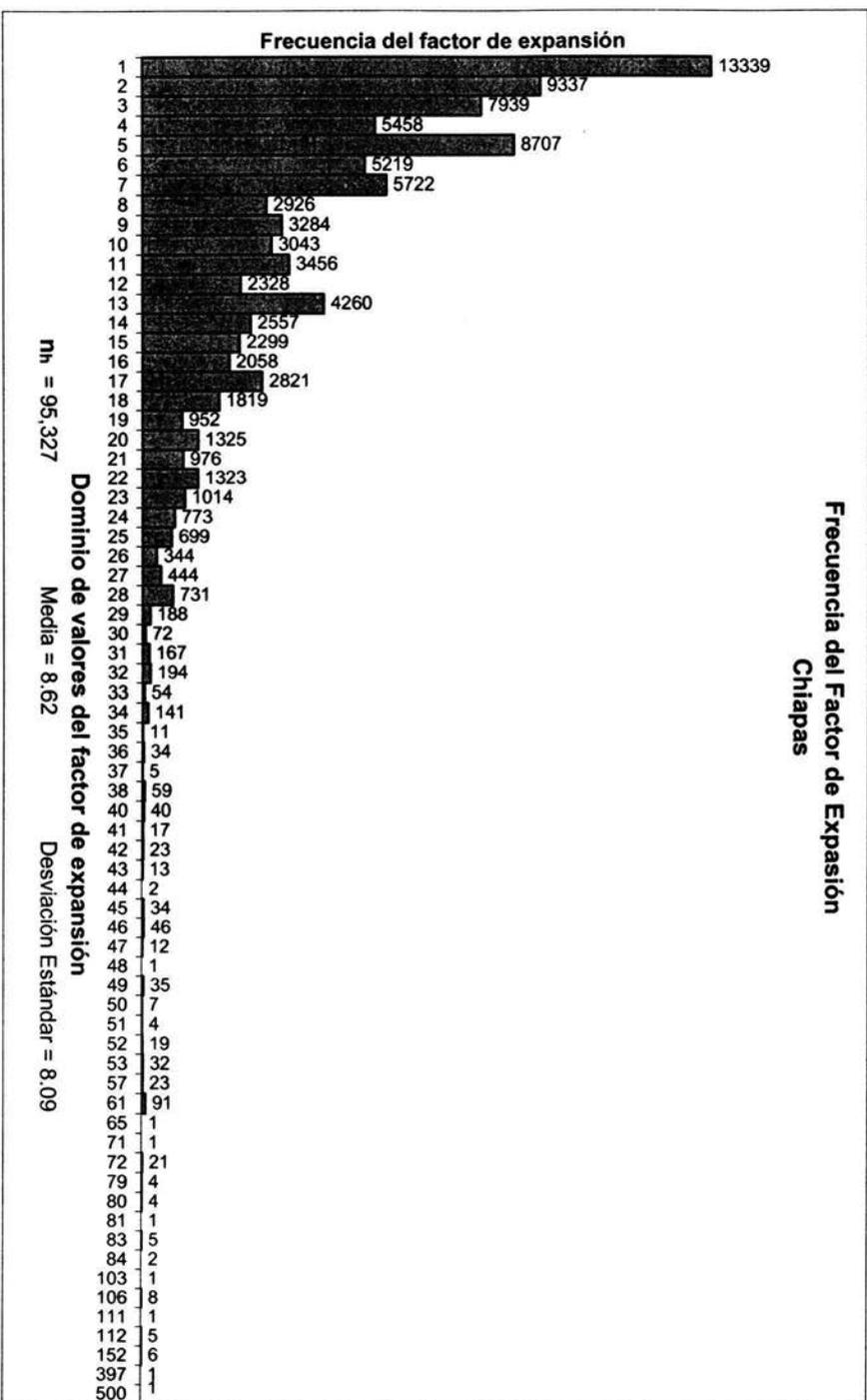


Figura A2.7 Frecuencia de los valores tomados por el Factor de expansión en Chiapas

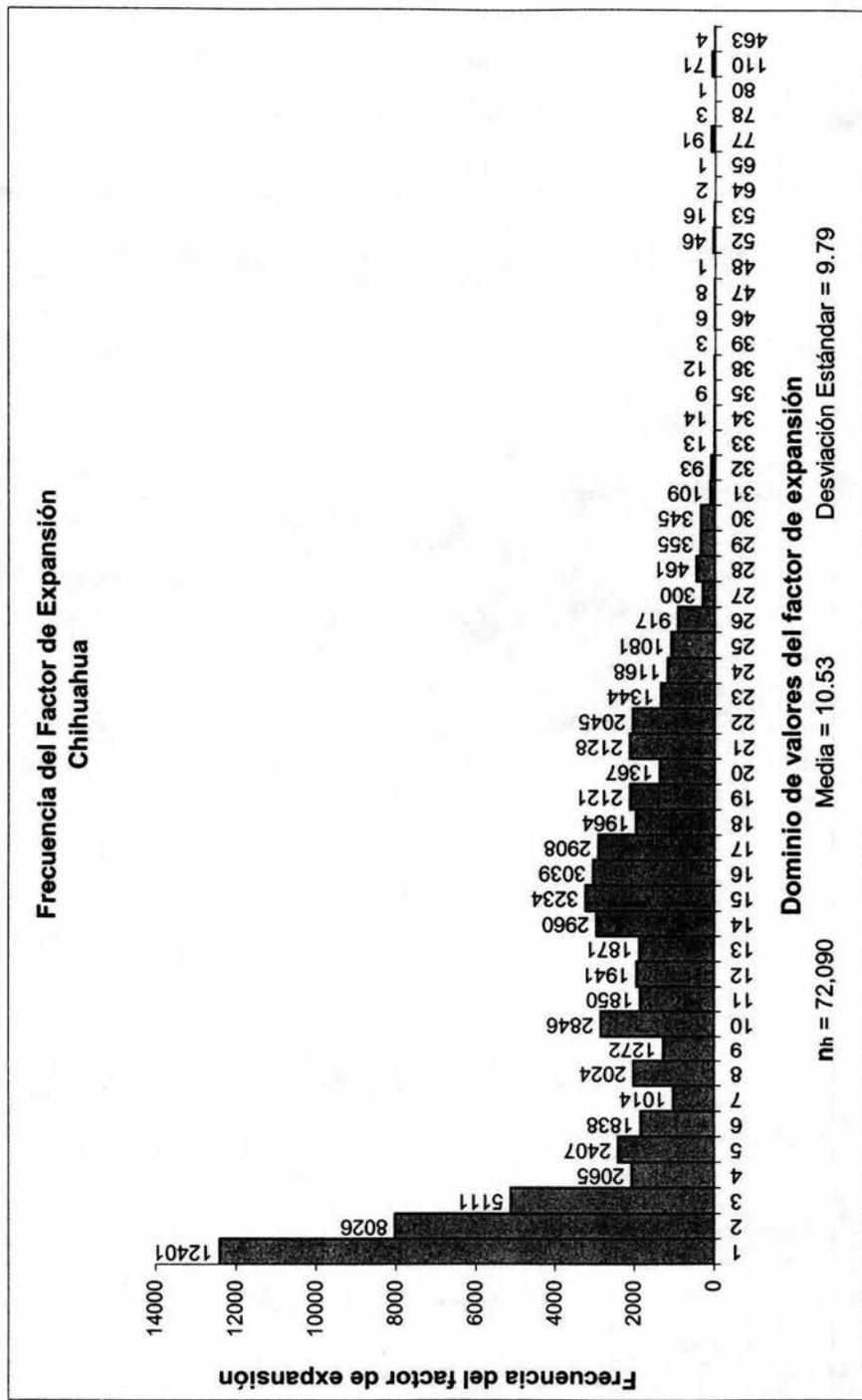


Figura A2.8 Frecuencia de los valores tomados por el Factor de expansión en Chihuahua

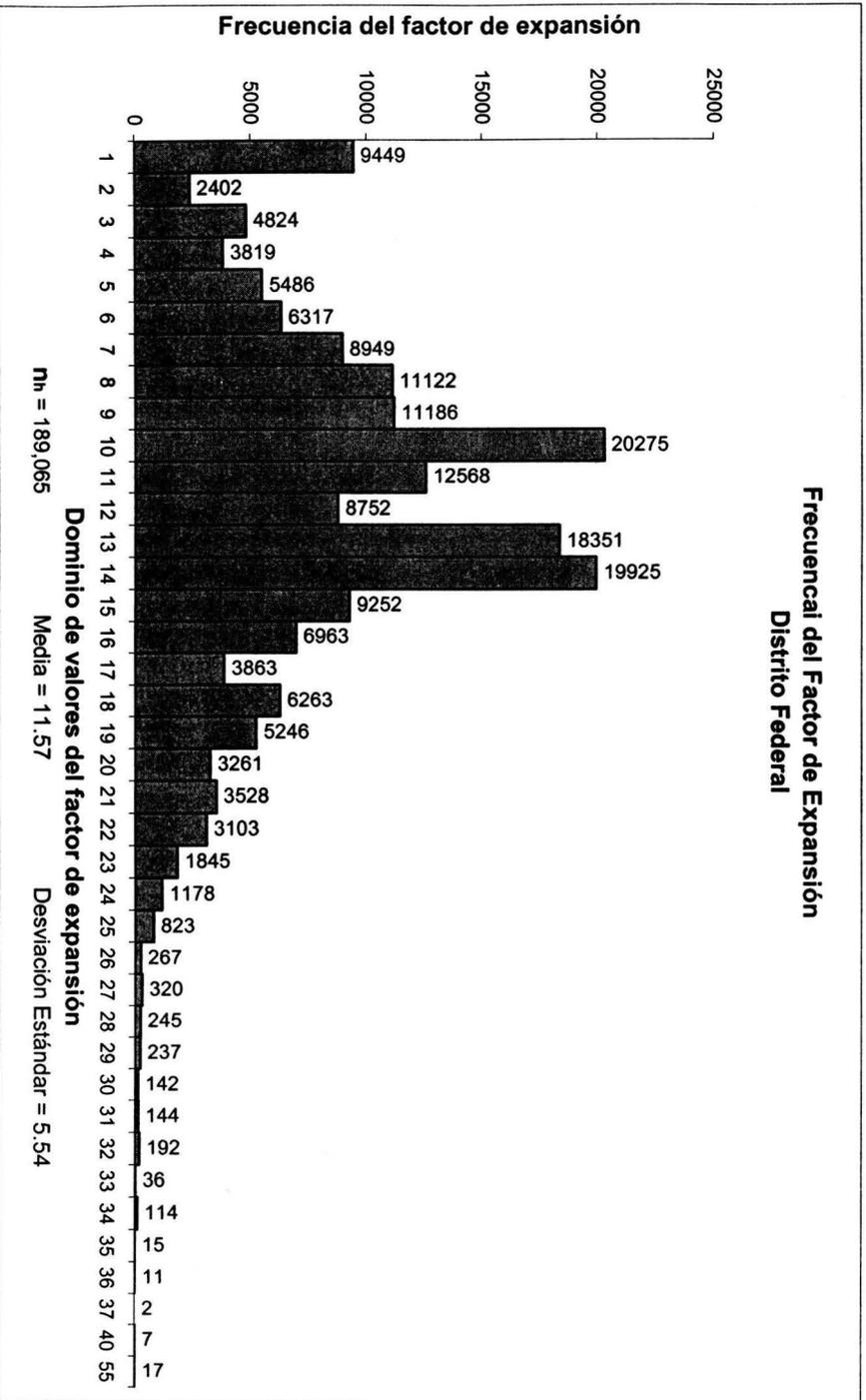


Figura A2.9 Frecuencia de los valores tomados por el Factor de expansión en el Distrito Federal

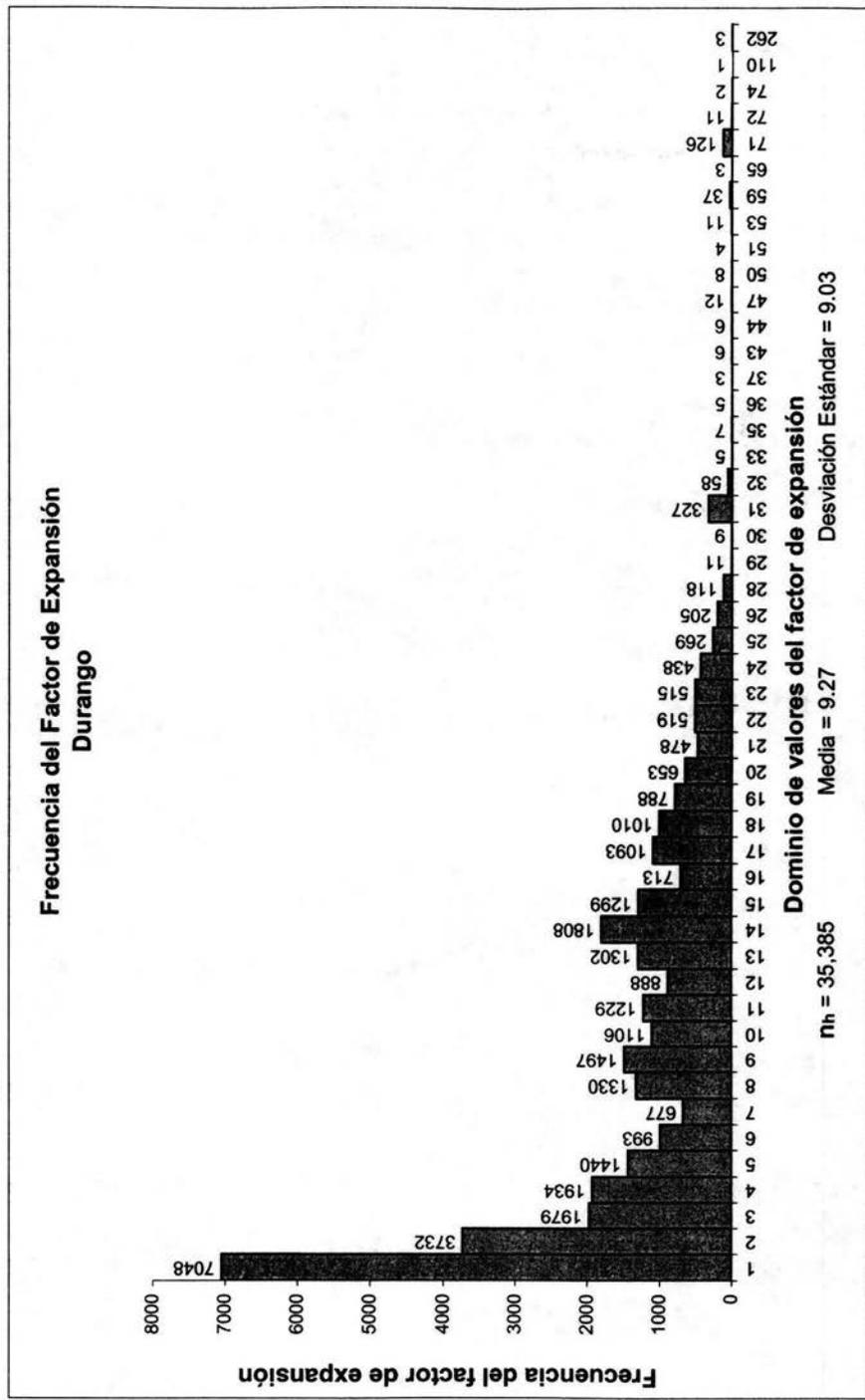


Figura A2.10 Frecuencia de los valores tomados por el Factor de expansión en Durango

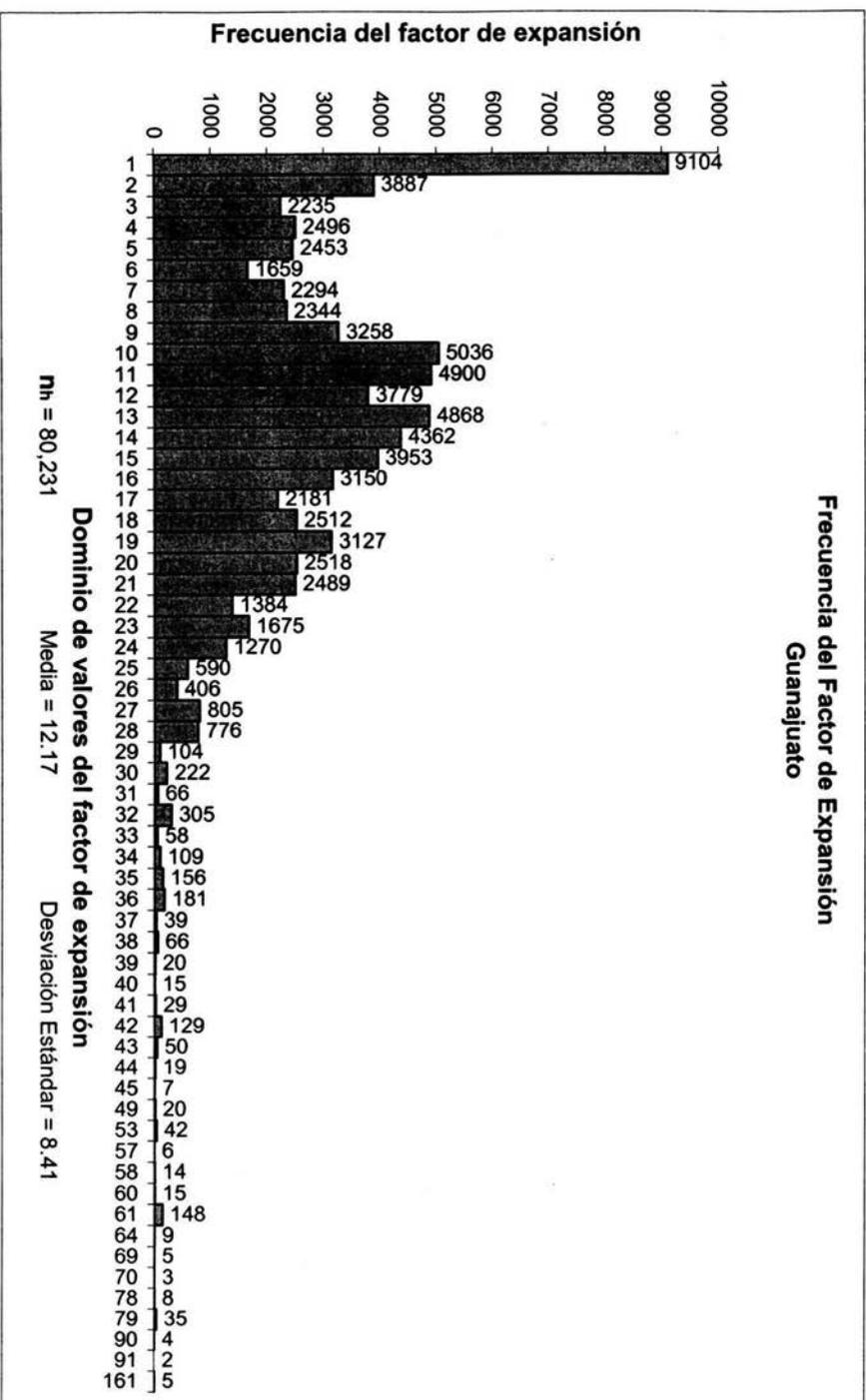


Figura A2.11 Frecuencia de los valores tomados por el Factor de expansión en Guanajuato

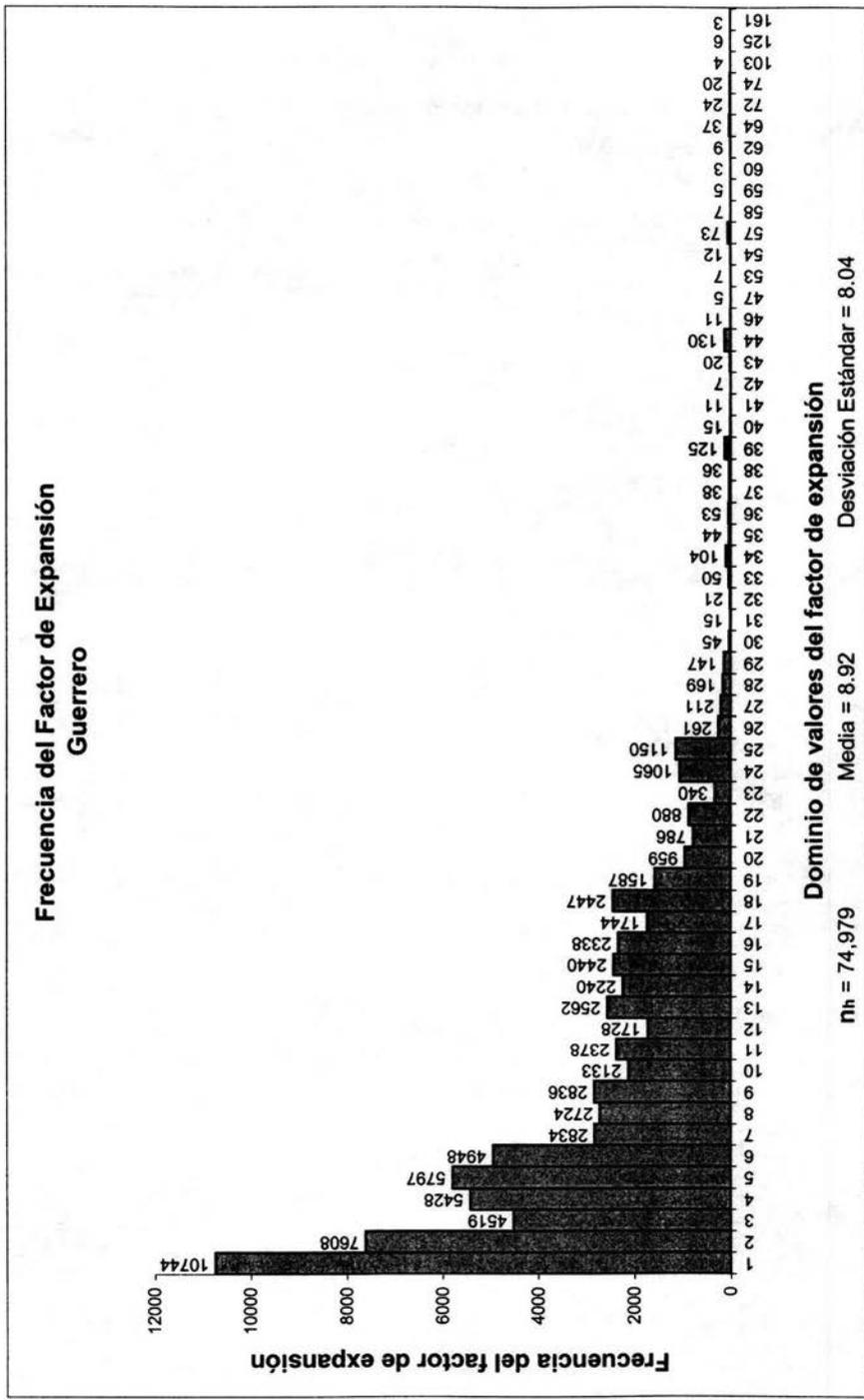


Figura A2.12 Frecuencia de los valores tomados por el Factor de expansión en Guerrero

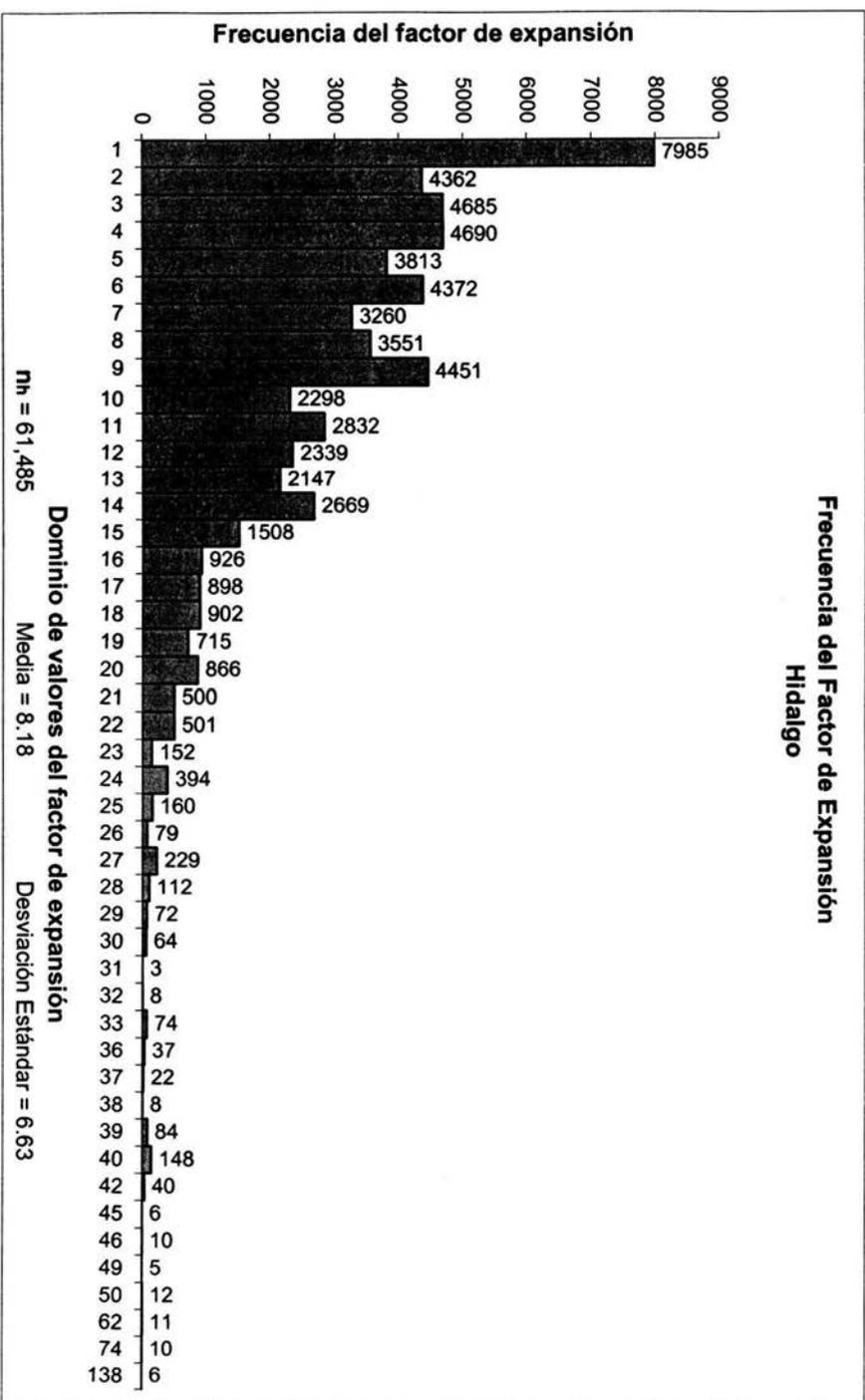


Figura A2.13 Frecuencia de los valores tomados por el Factor de expansión en Hidalgo

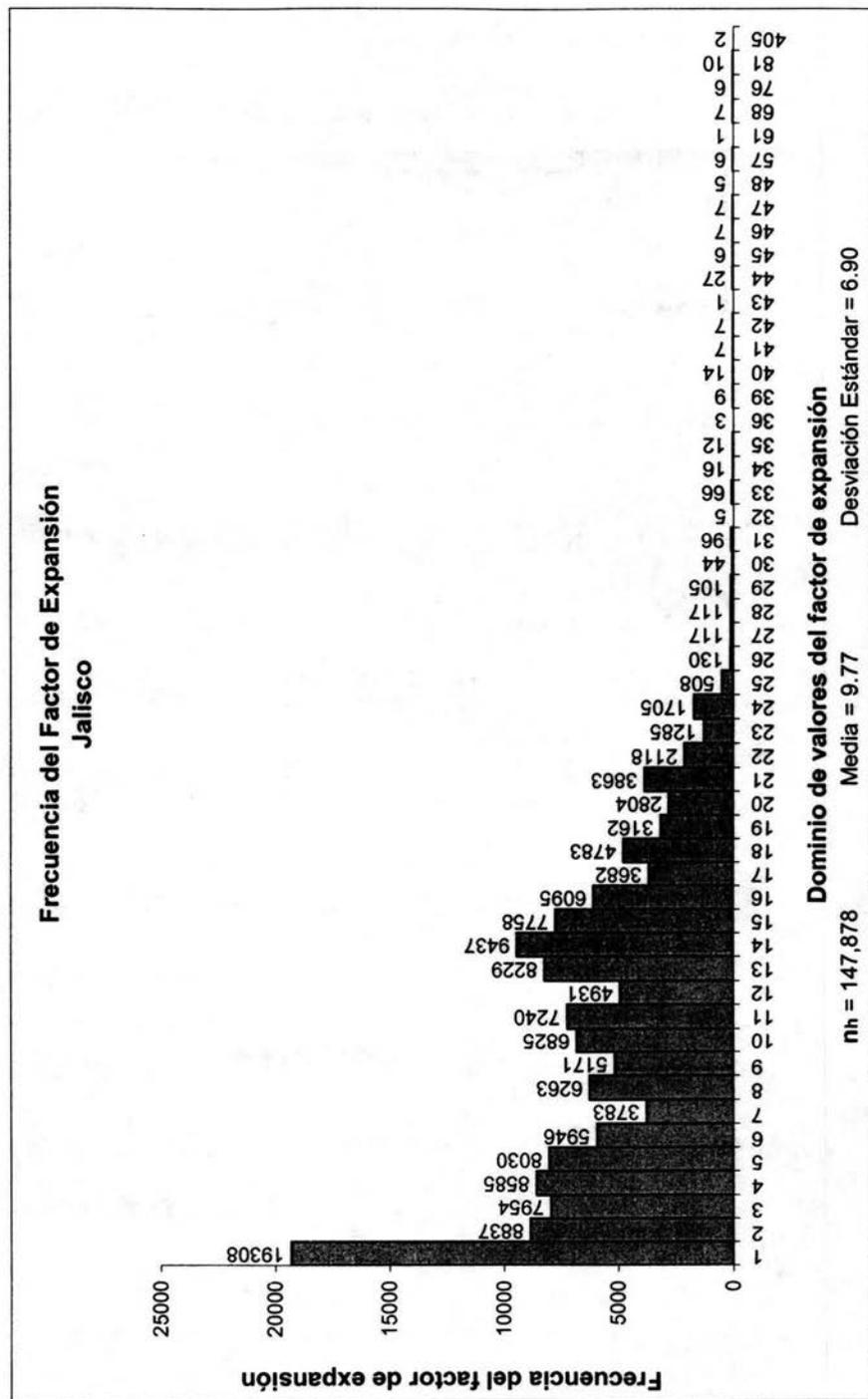


Figura A2.14 Frecuencia de los valores tomados por el Factor de expansión en Jalisco

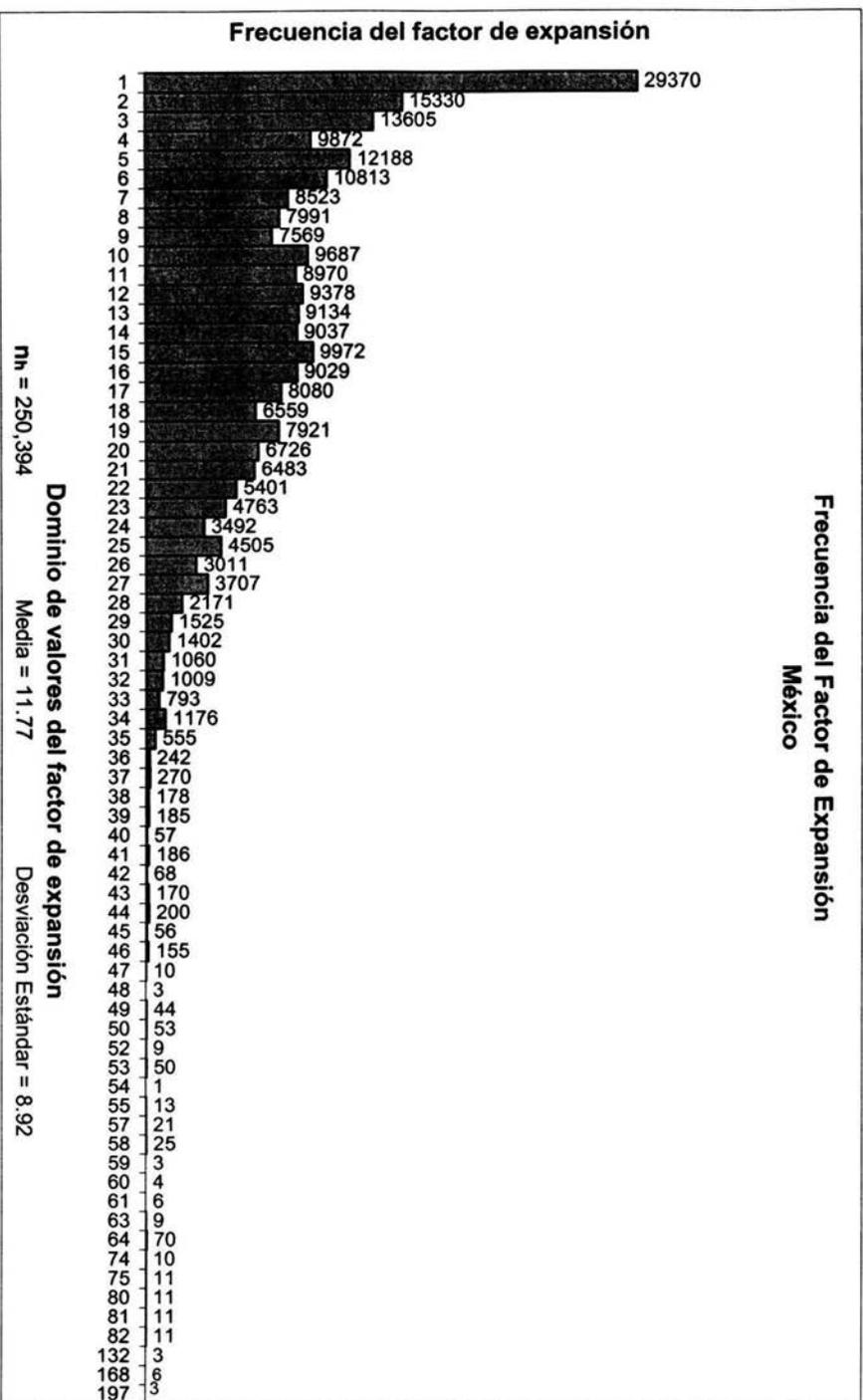


Figura A2.15 Frecuencia de los valores tomados por el Factor de expansión en el Edo. de México

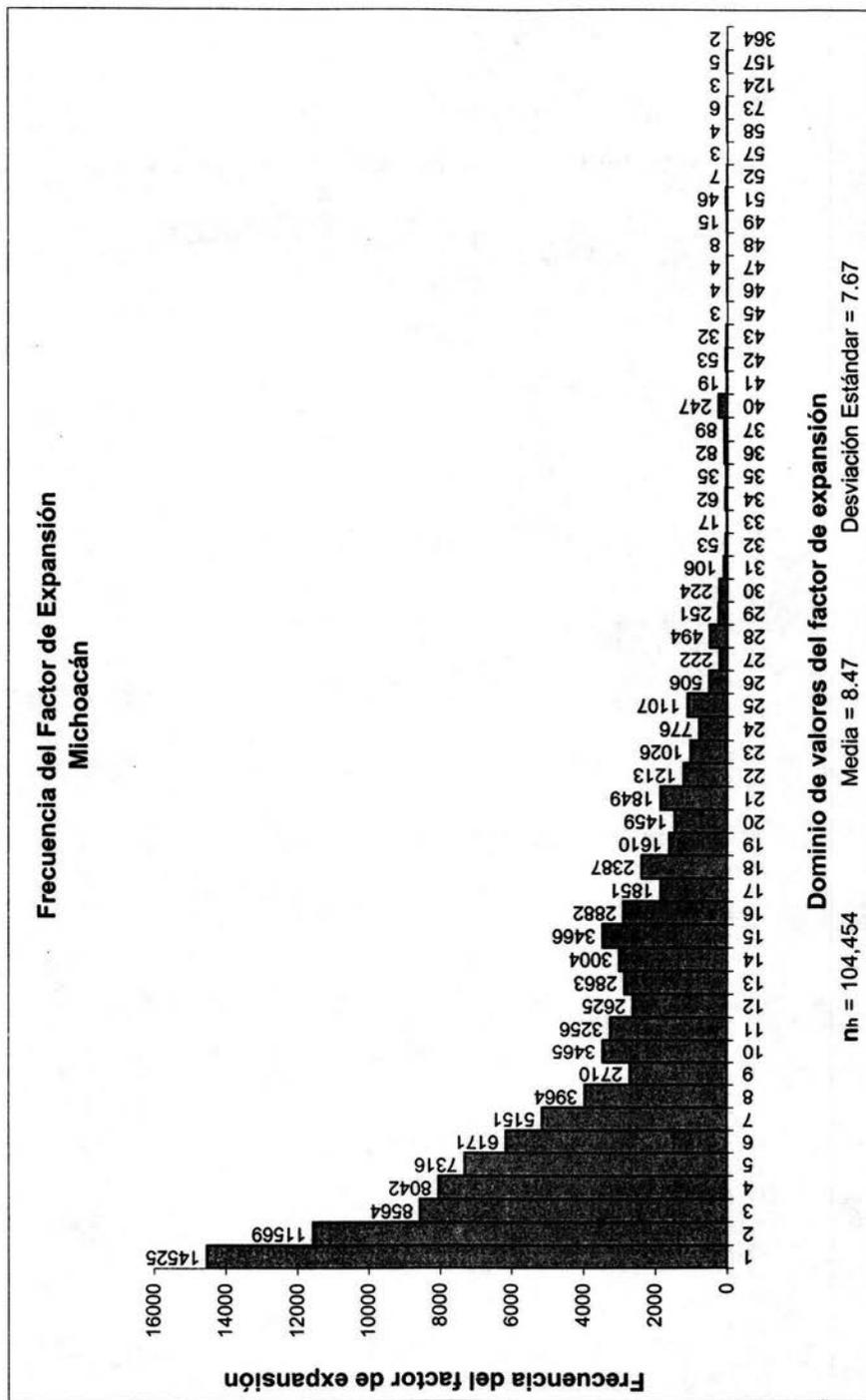


Figura A2.16 Frecuencia de los valores tomados por el Factor de expansión en Michoacán

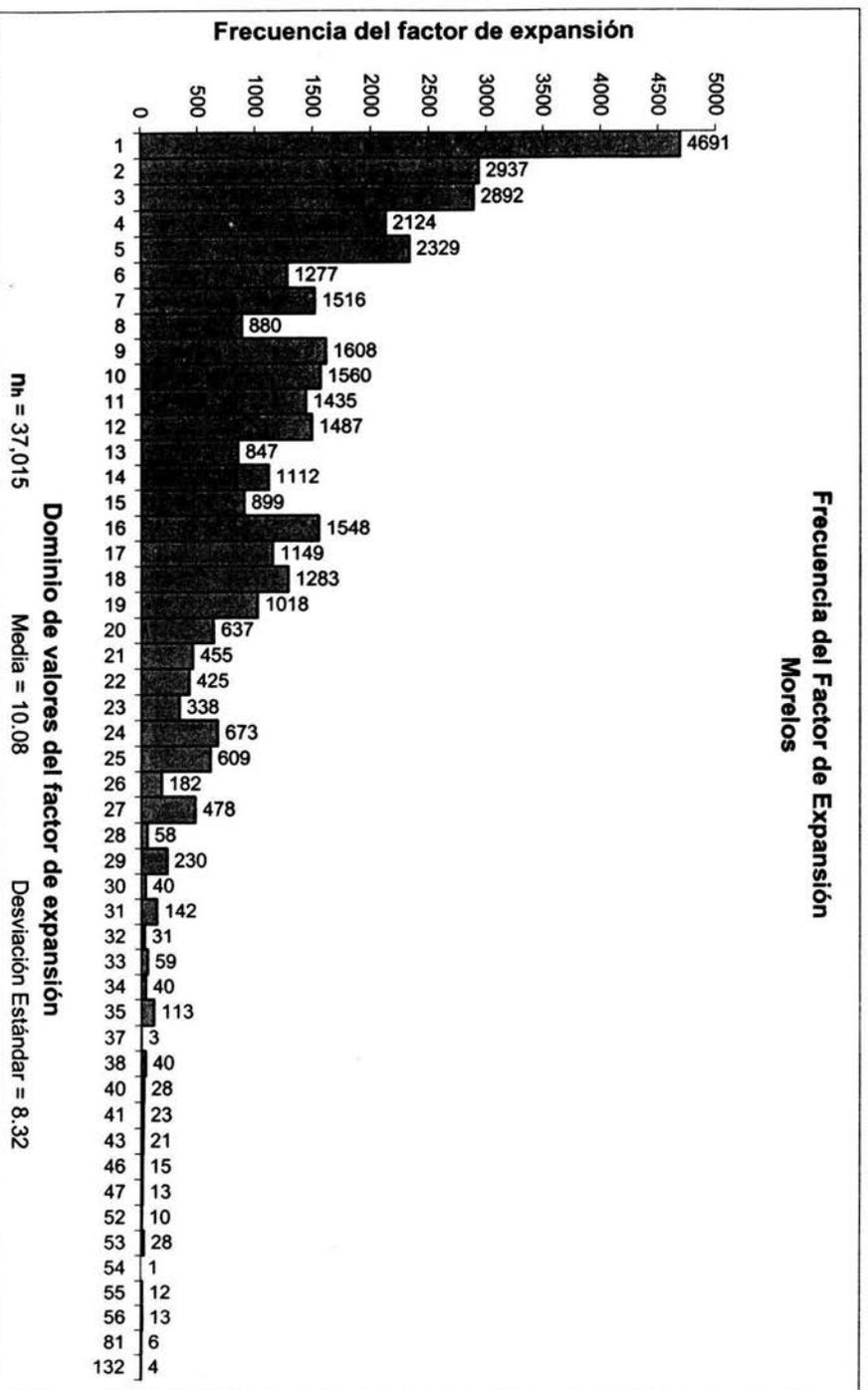


Figura A2.17 Frecuencia de los valores tomados por el Factor de expansión en Morelos

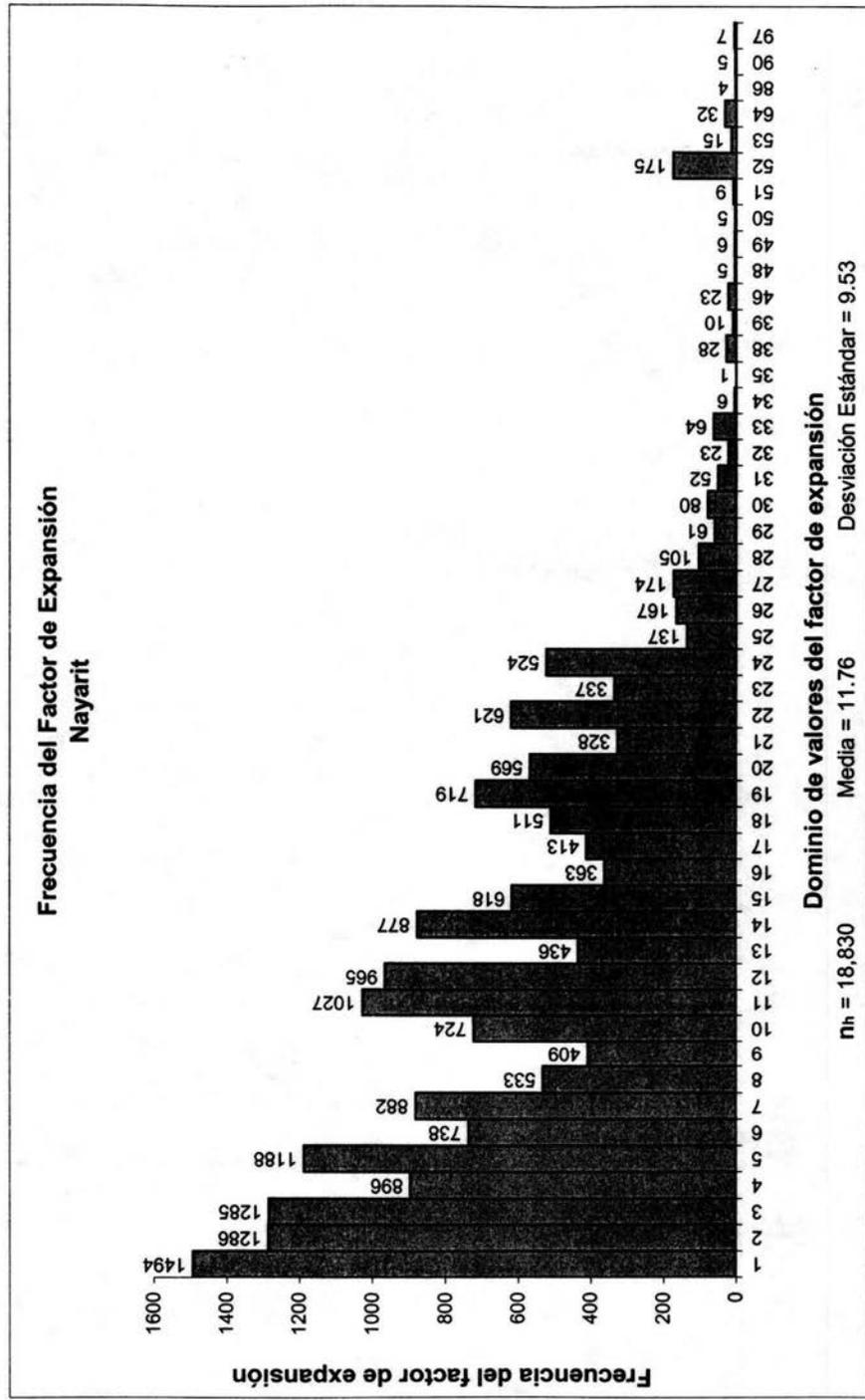


Figura A2.18 Frecuencia de los valores tomados por el Factor de expansión en Nayarit

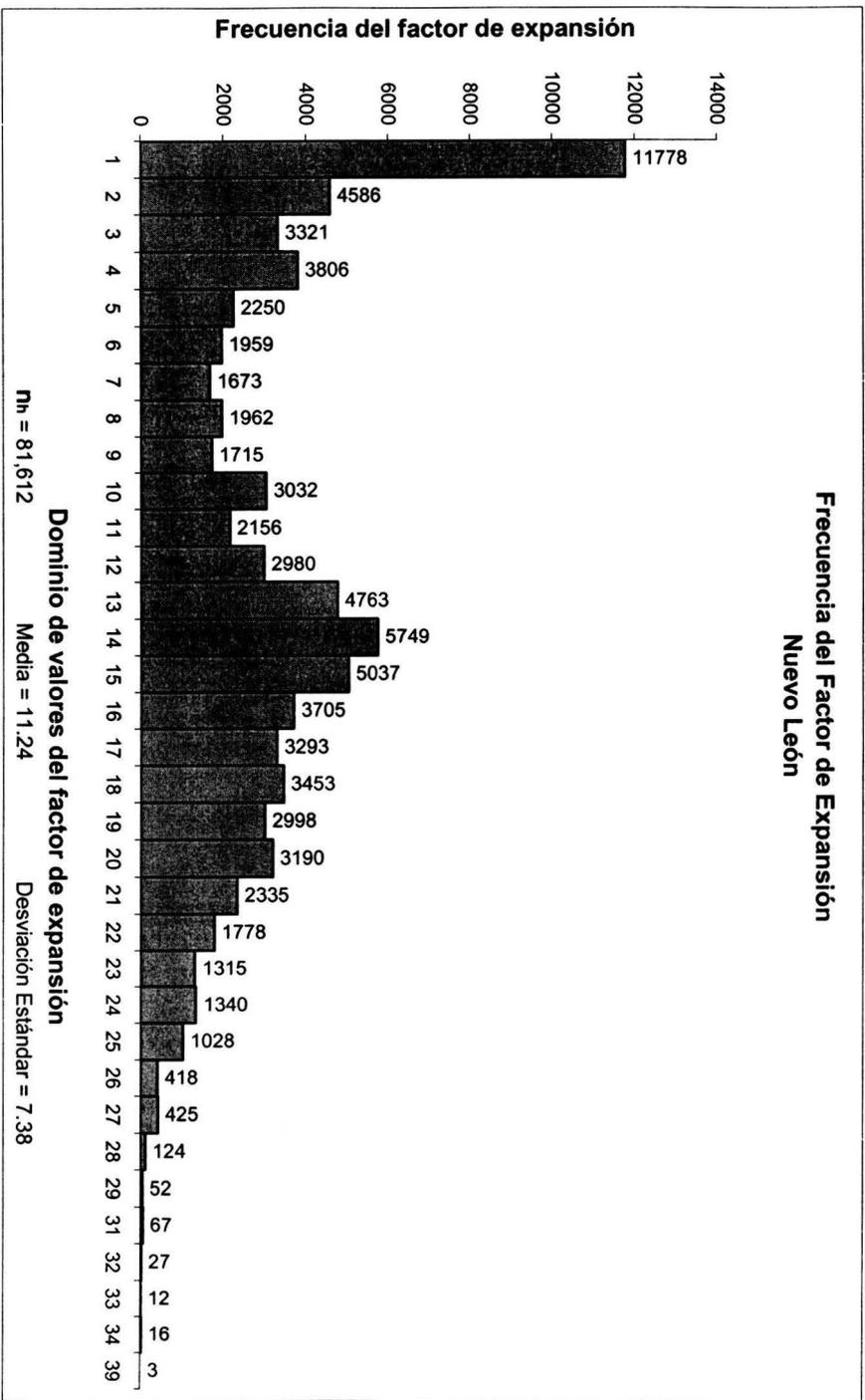


Figura A2.19 Frecuencia de los valores tomados por el Factor de expansión en Nuevo León

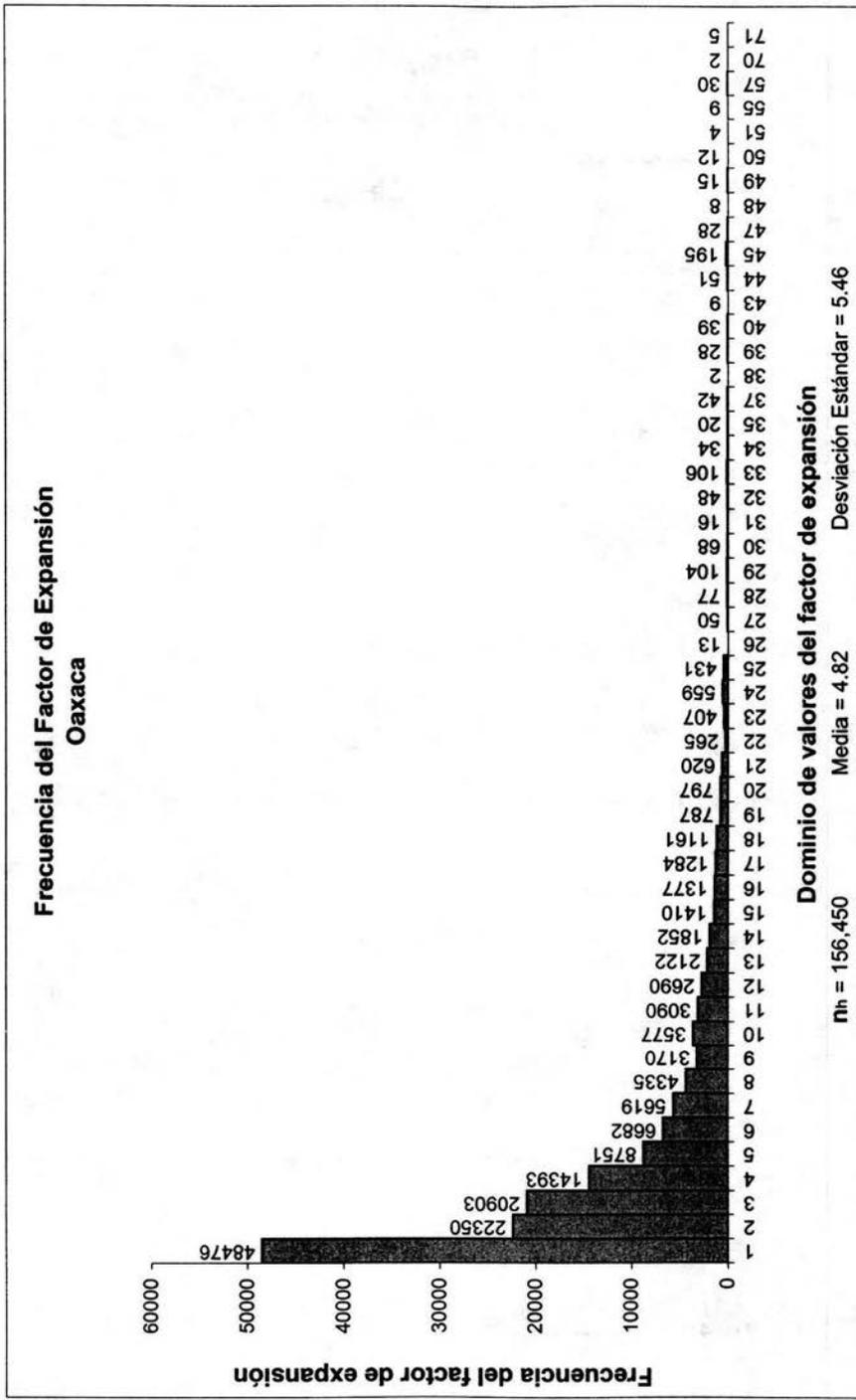


Figura A2.20 Frecuencia de los valores tomados por el Factor de expansión en Oaxaca

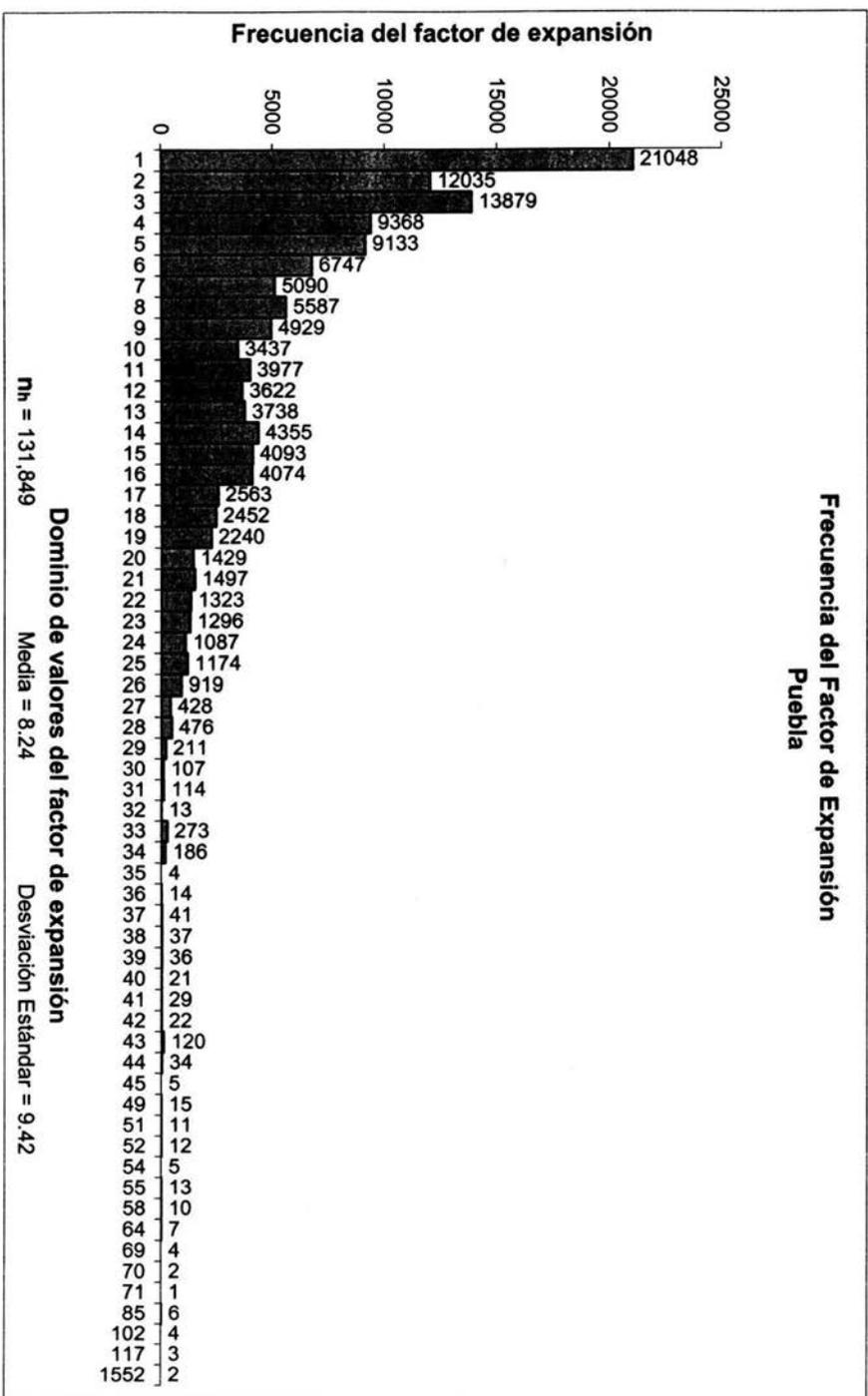


Figura A2.21 Frecuencia de los valores tomados por el Factor de expansión en Puebla

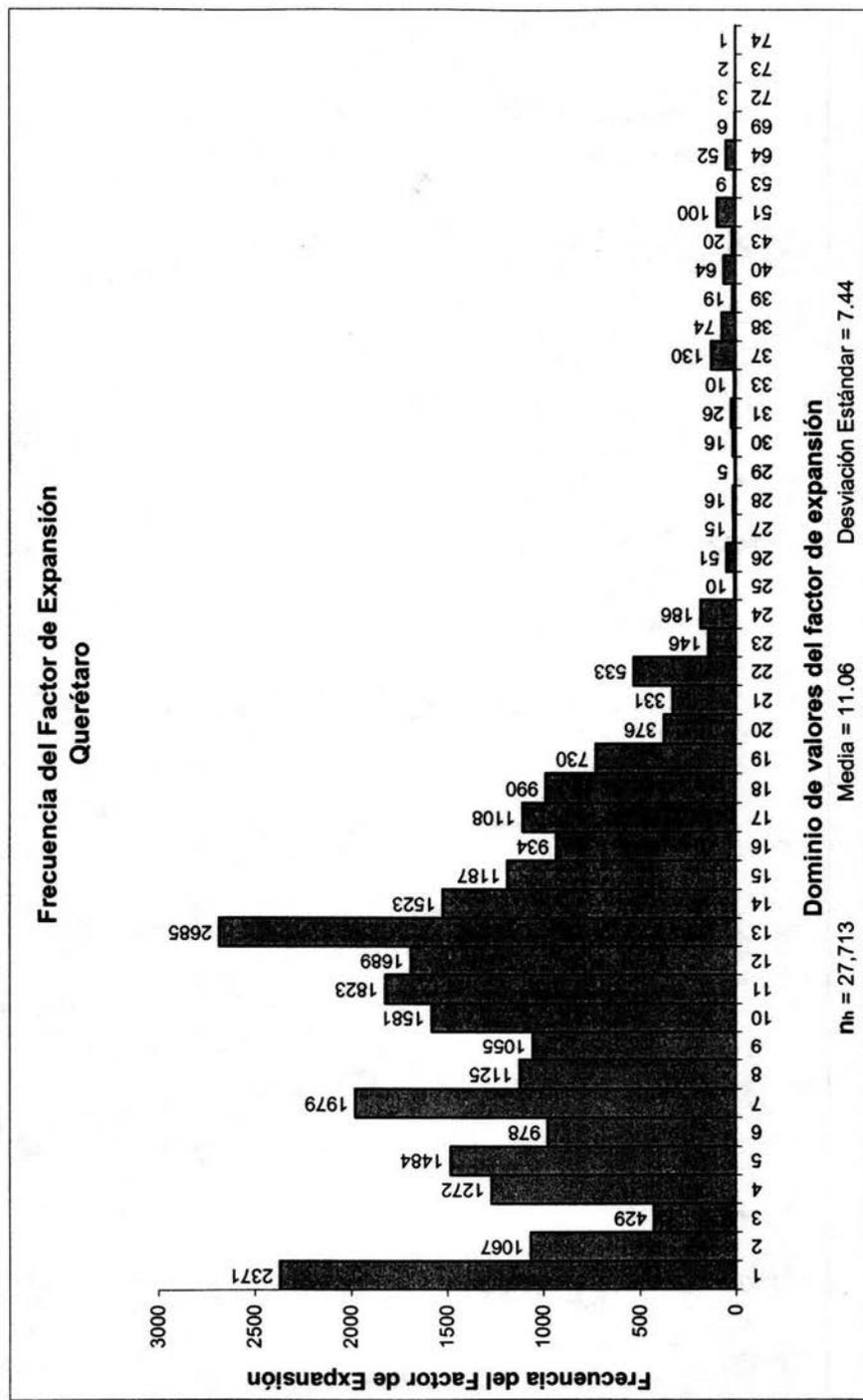


Figura A2.22 Frecuencia de los valores tomados por el Factor de expansión en Querétaro

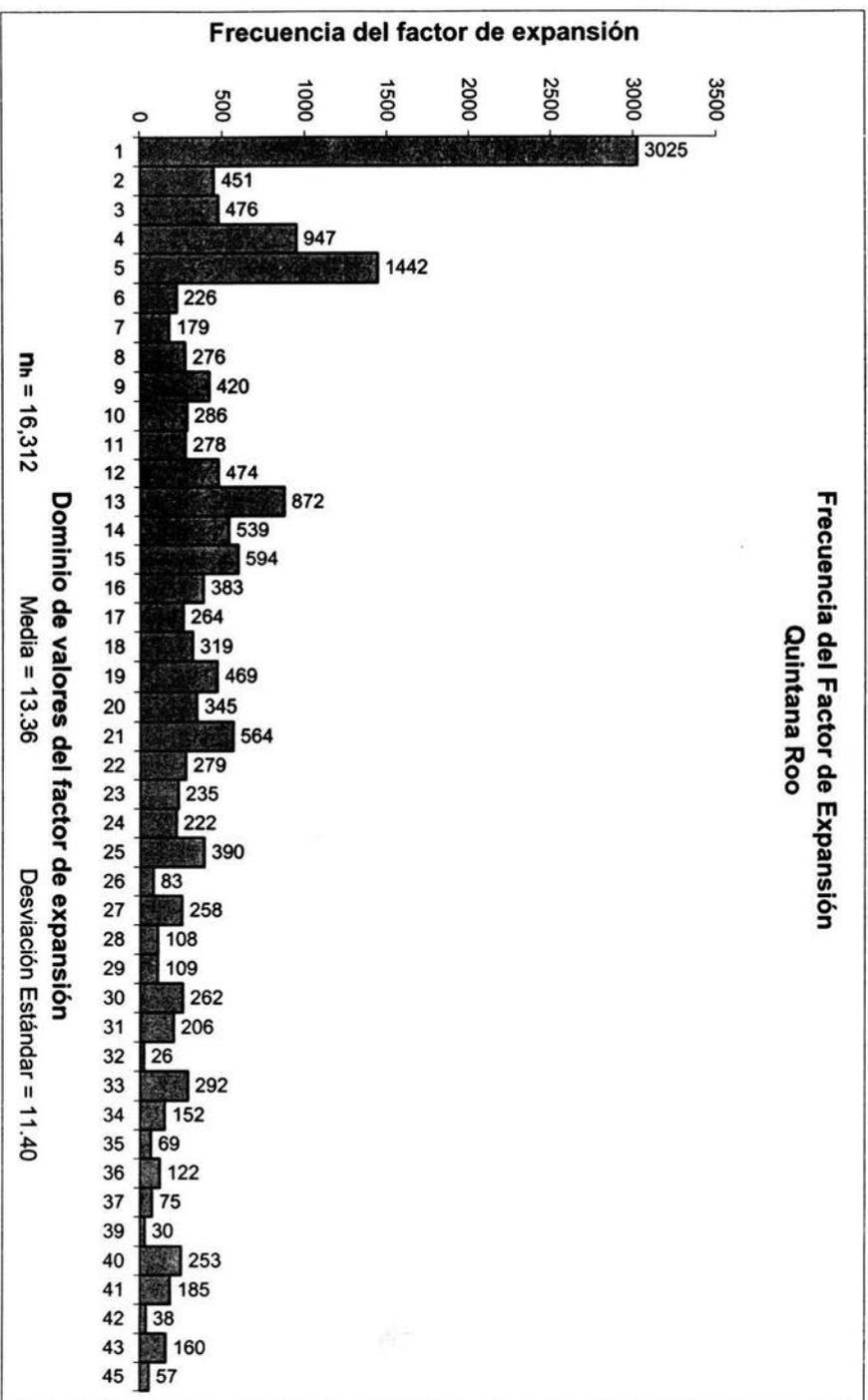


Figura A2.23 Frecuencia de los valores tomados por el Factor de expansión en Quintana Roo

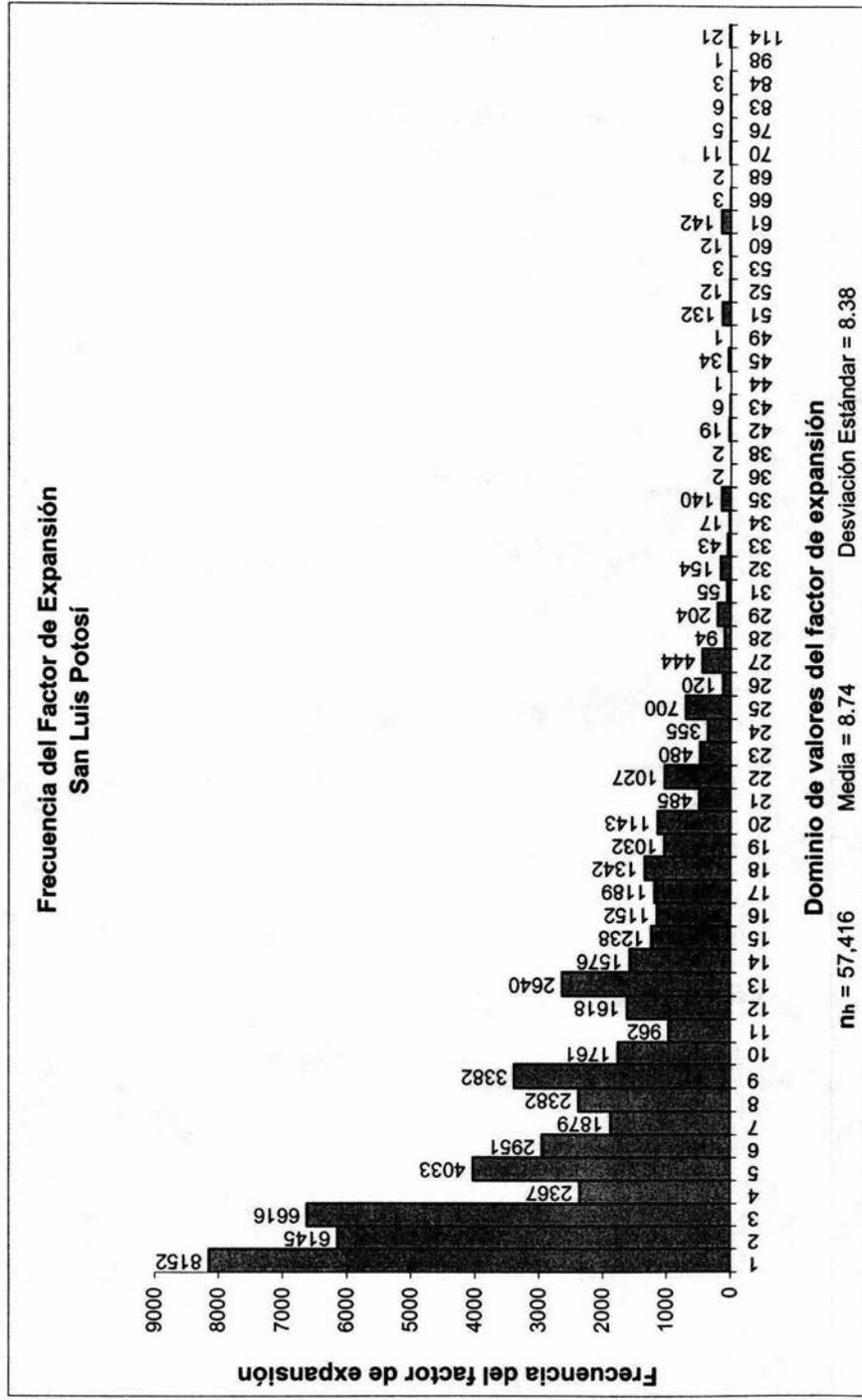


Figura A2.24 Frecuencia de los valores tomados por el Factor de expansión en San Luis Potosí

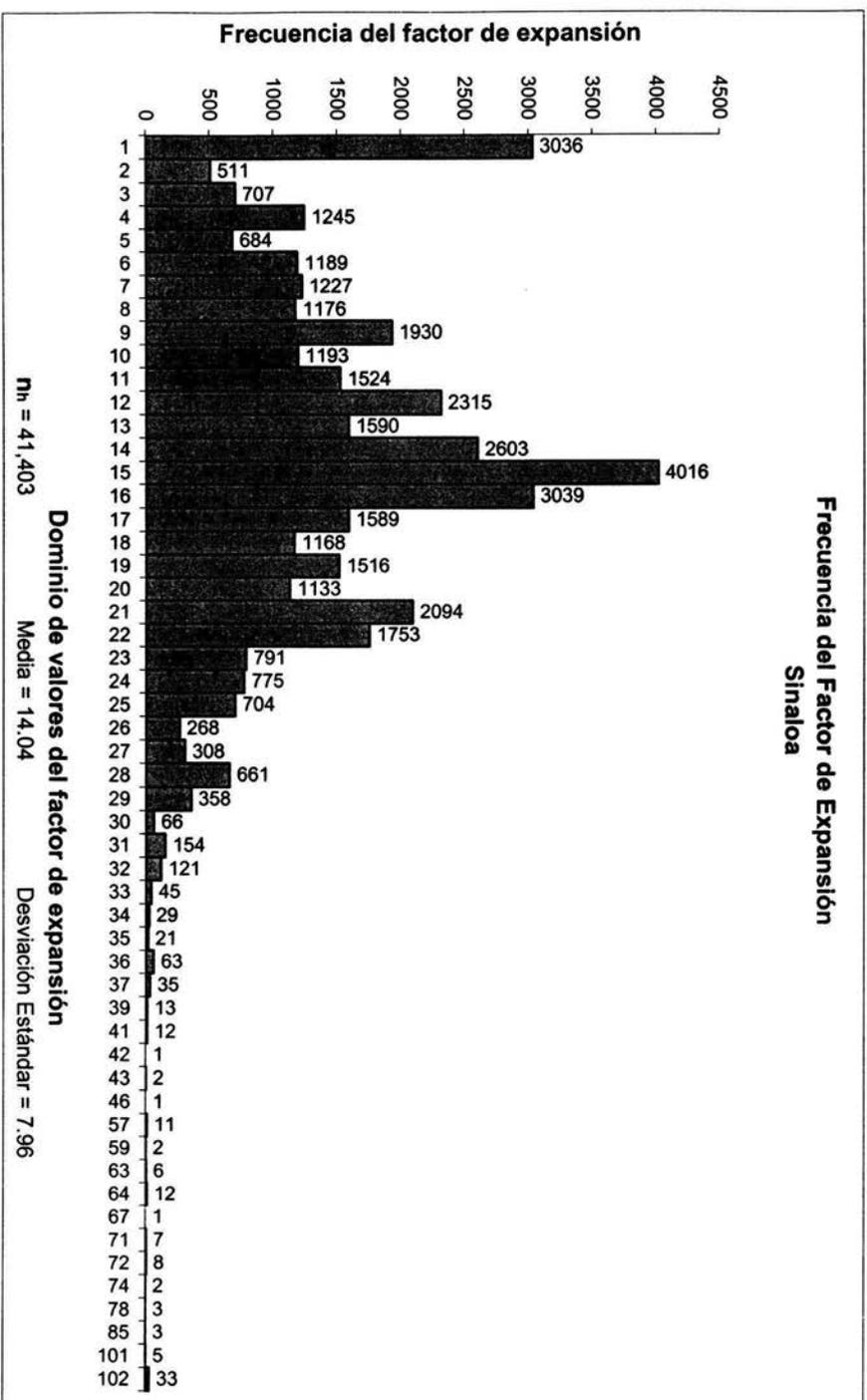


Figura A2.25 Frecuencia de los valores tomados por el Factor de expansión en Sinaloa

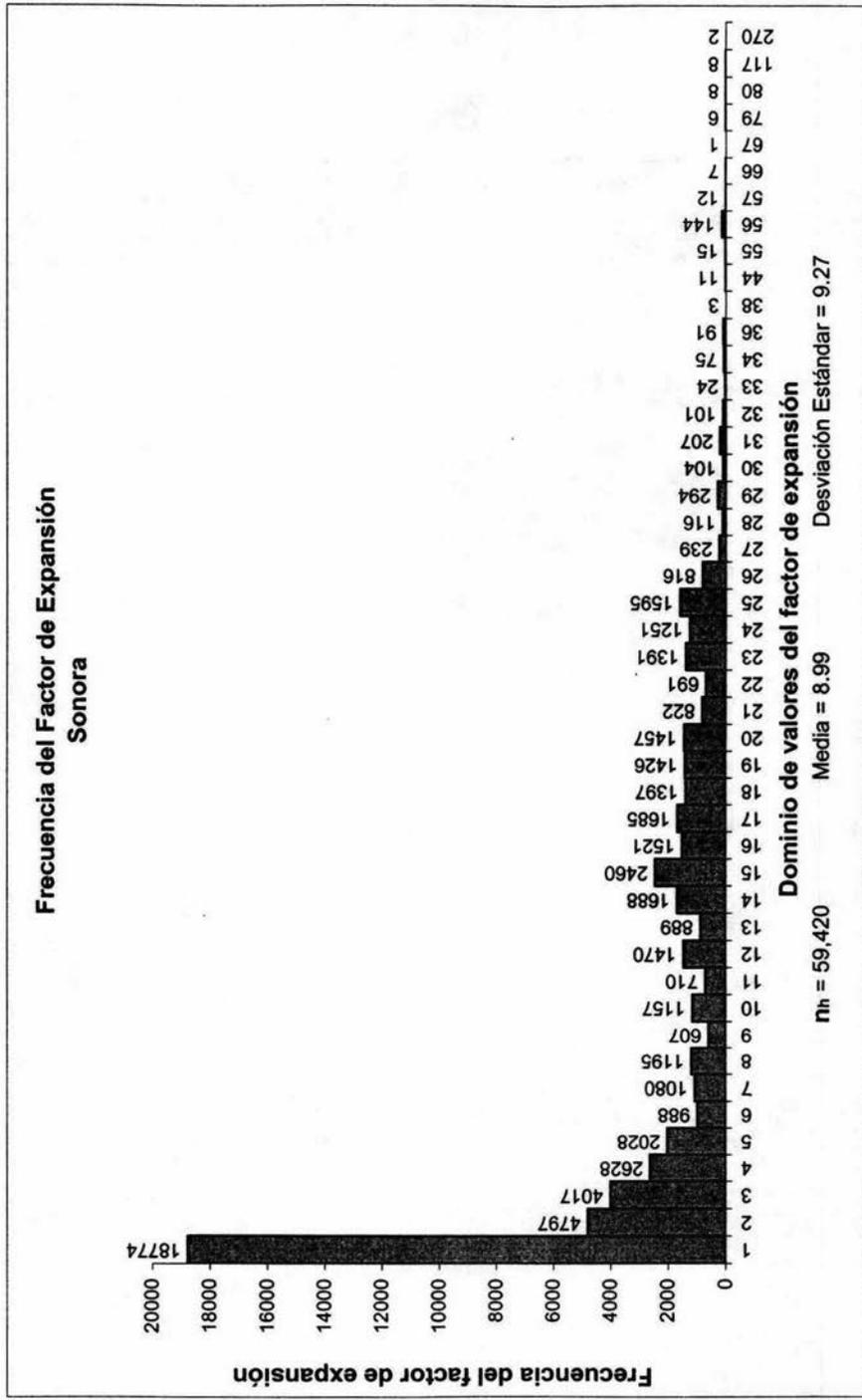


Figura A2.26 Frecuencia de los valores tomados por el Factor de expansión en Sonora

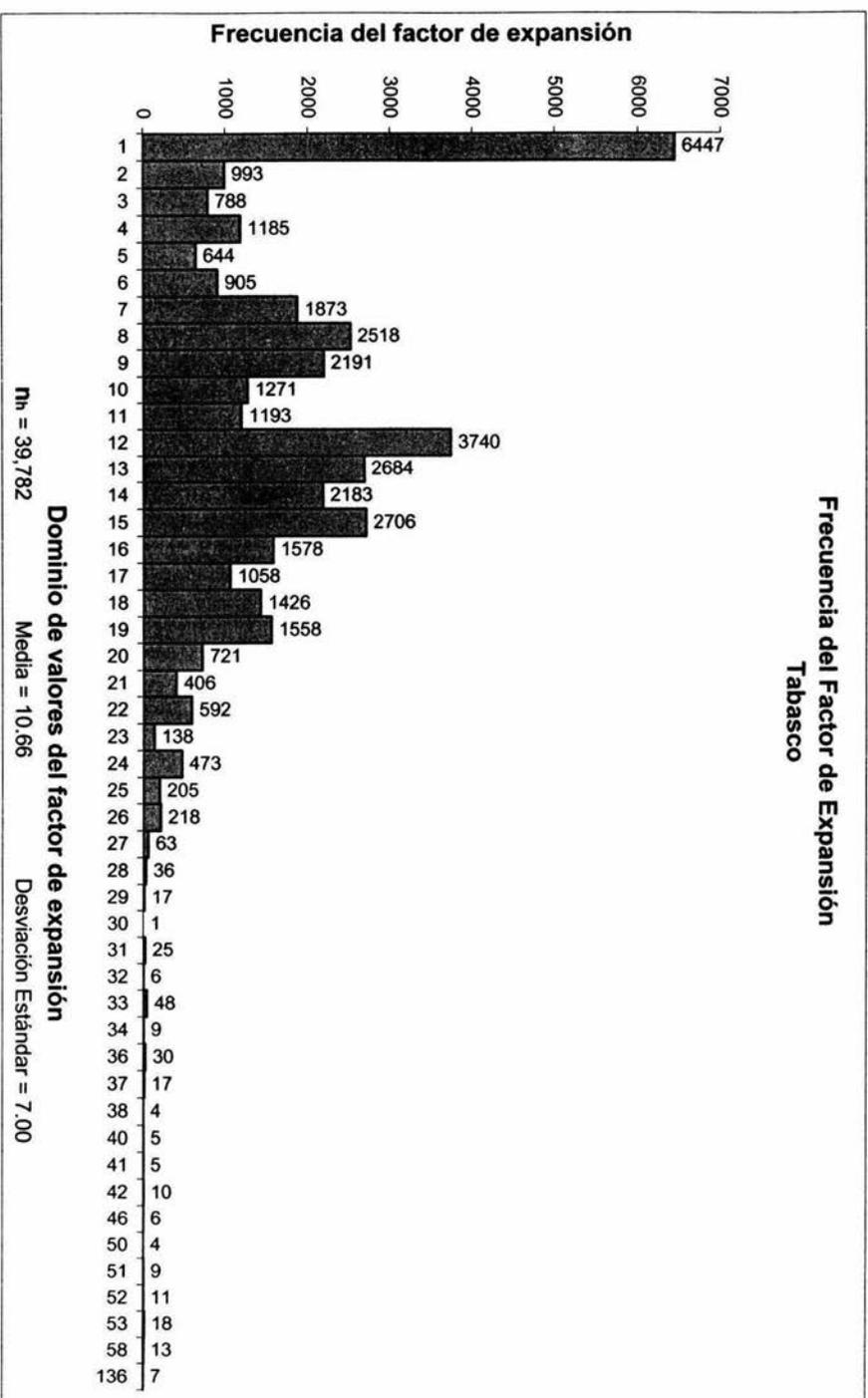


Figura A2.27 Frecuencia de los valores tomados por el Factor de expansión en Tabasco

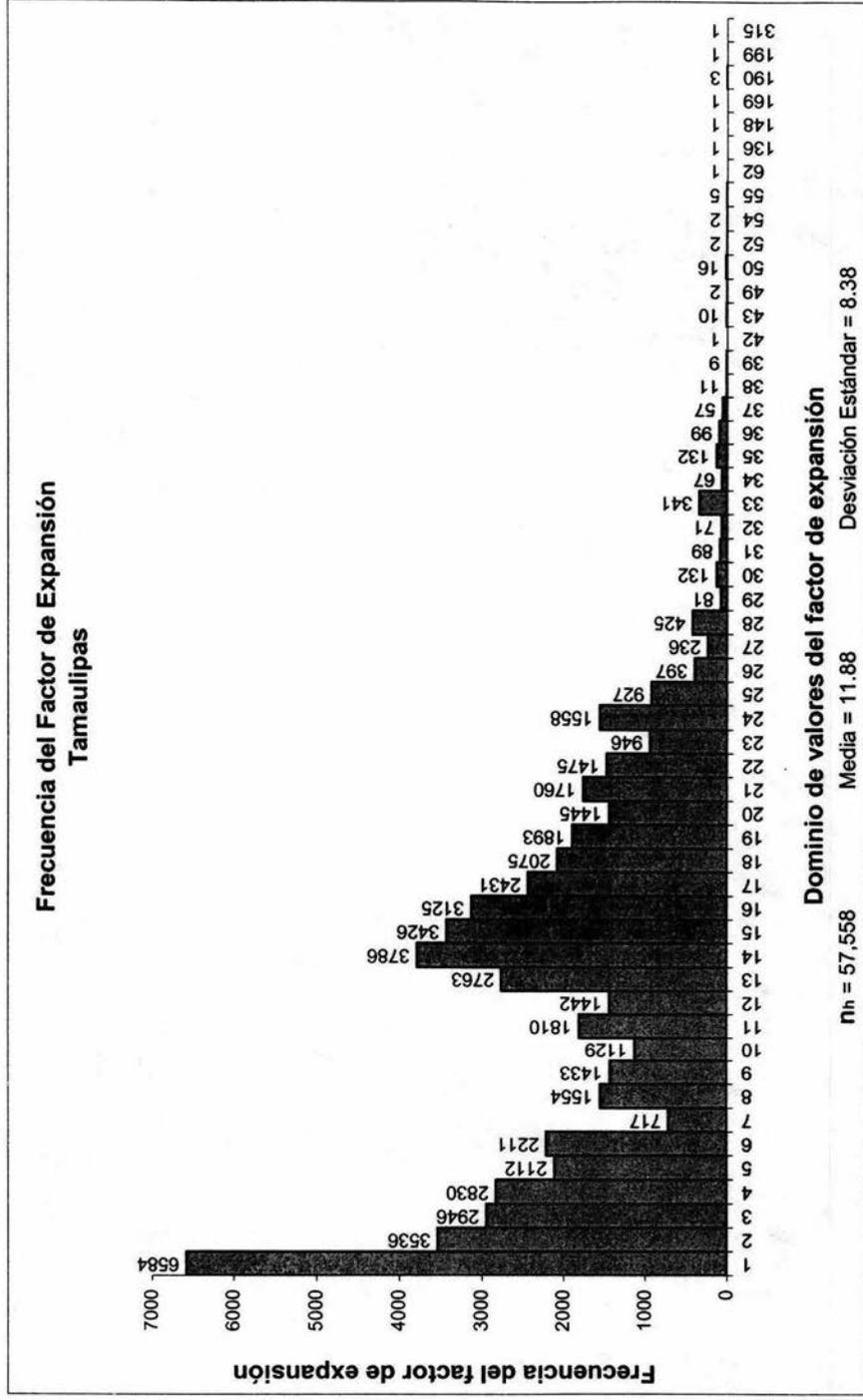


Figura A2.28 Frecuencia de los valores tomados por el Factor de expansión en Tamaulipas

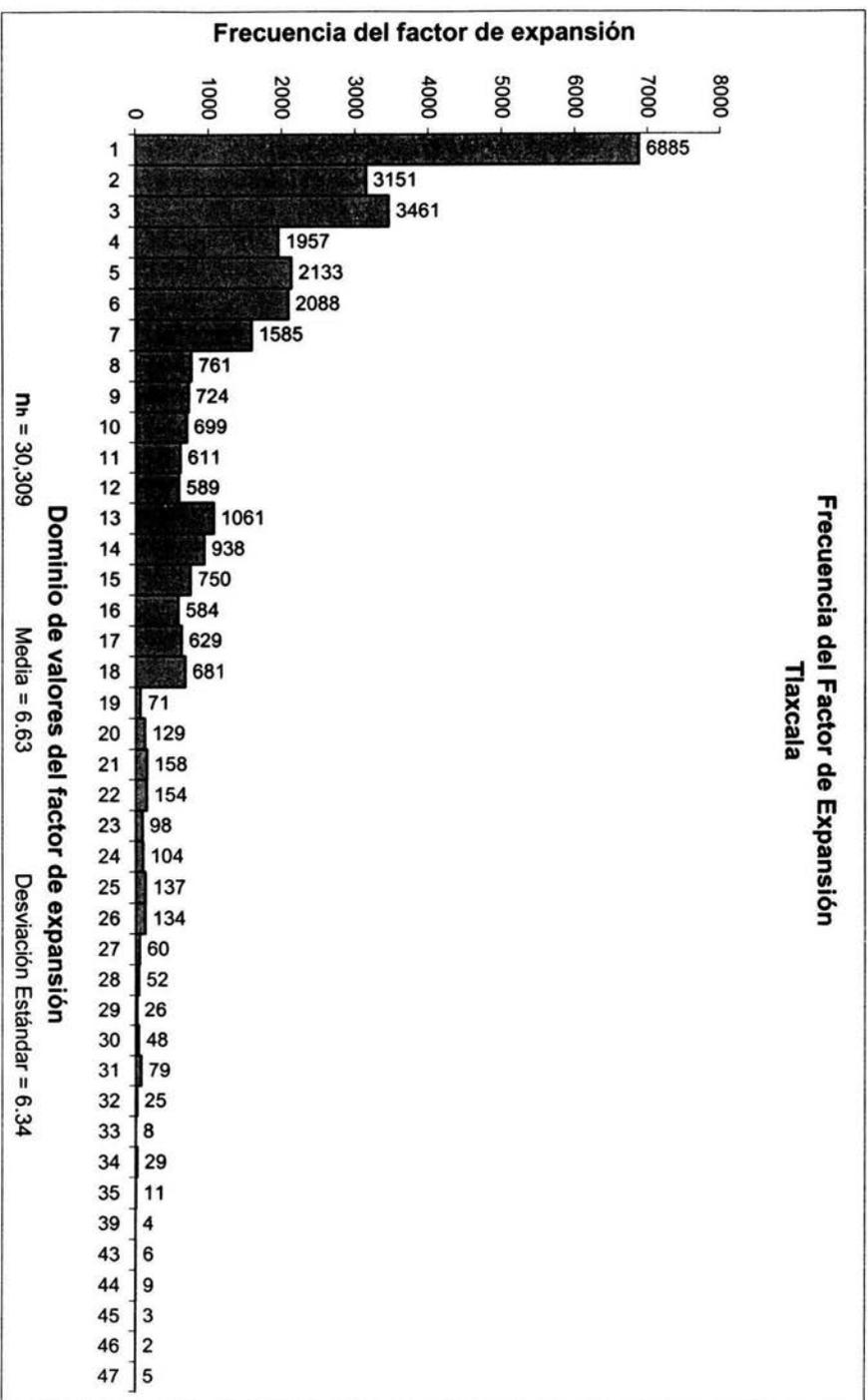


Figura A2.29 Frecuencia de los valores tomados por el Factor de expansión en Tlaxcala

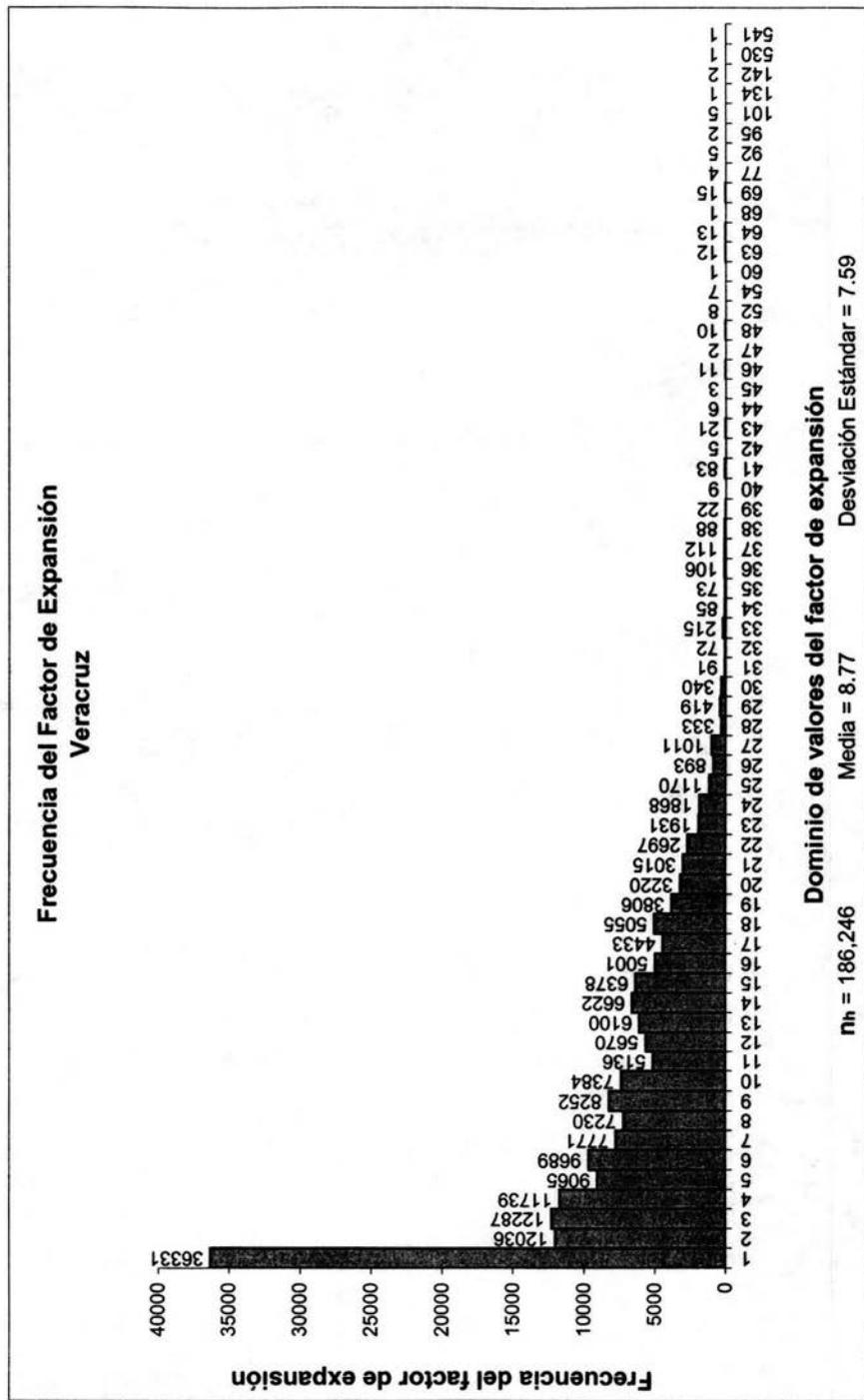


Figura A2.30 Frecuencia de los valores tomados por el Factor de expansión en Veracruz

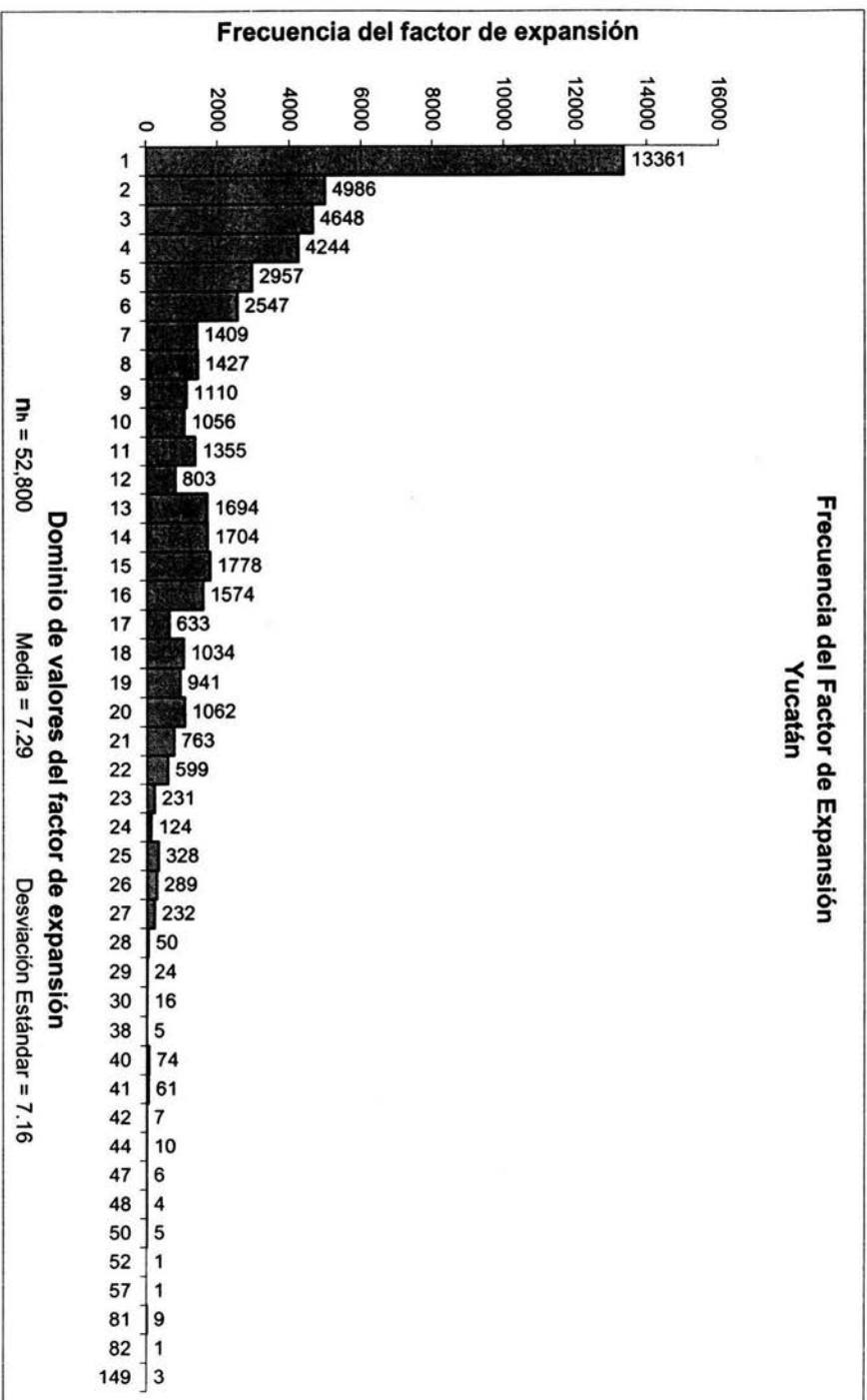


Figura A2.31 Frecuencia de los valores tomados por el Factor de expansión en Yucatán

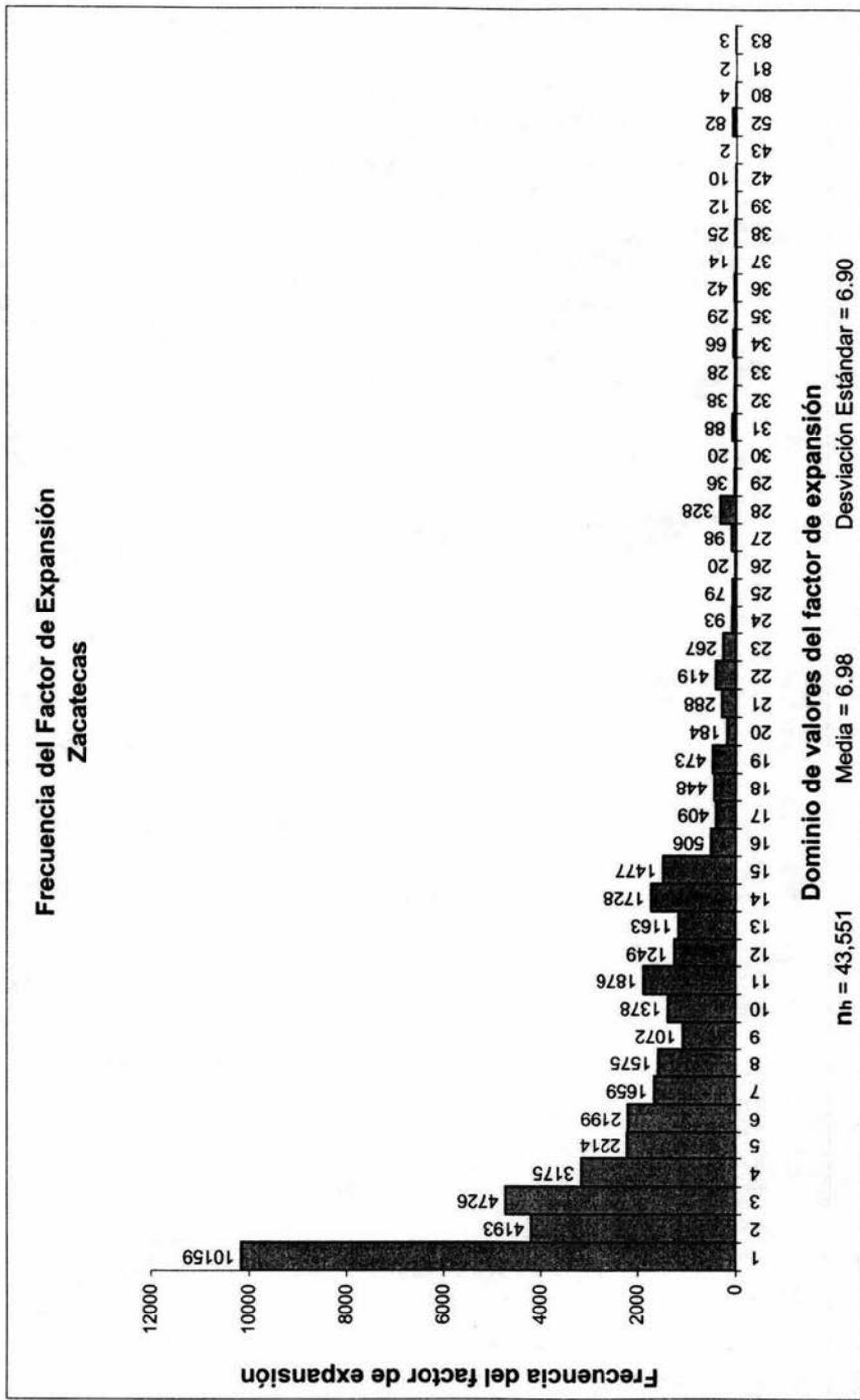


Figura A2.32 Frecuencia de los valores tomados por el Factor de expansión en Zacatecas

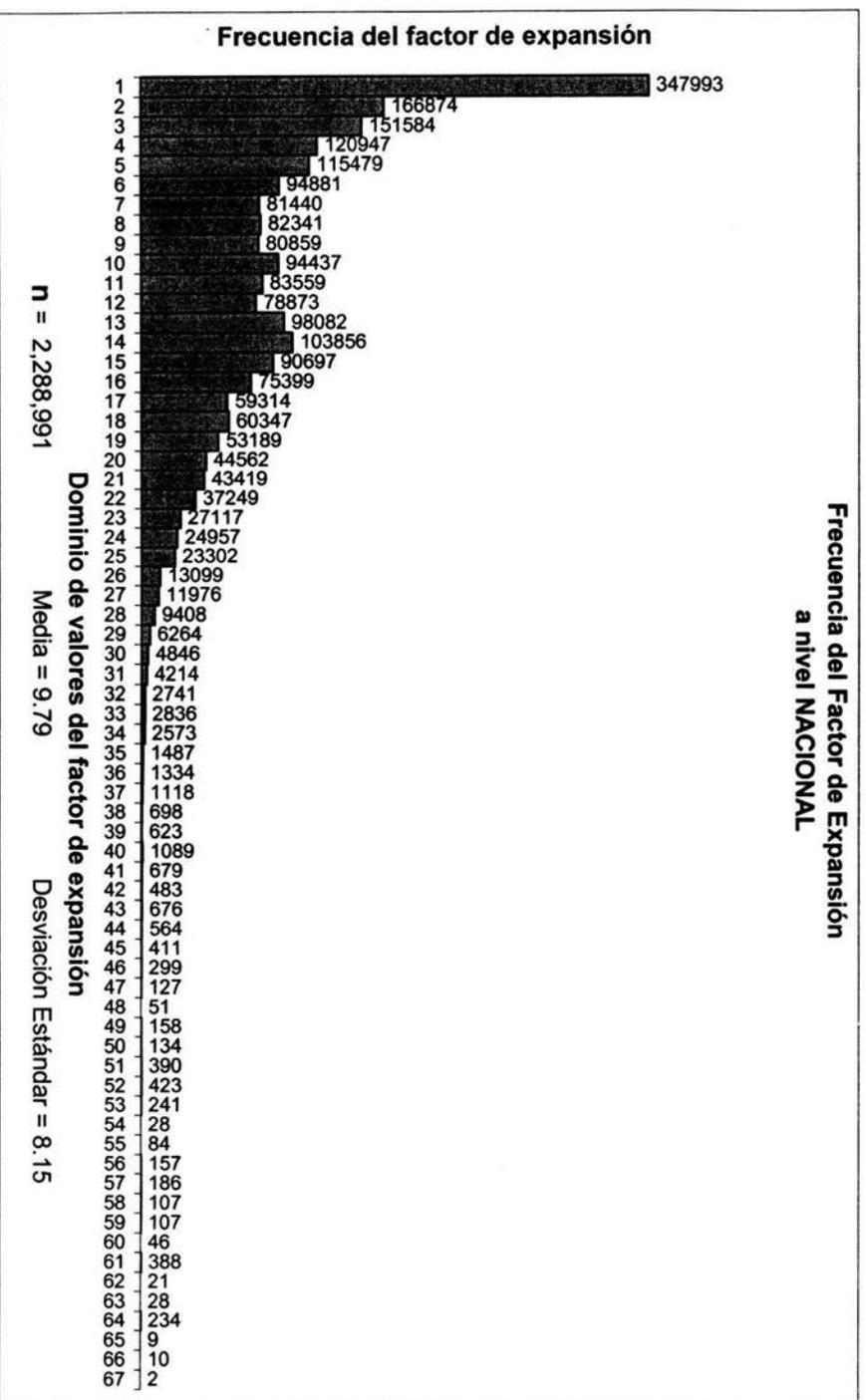


Figura A2.33 Frecuencia de los valores tomados por el Factor de expansión a nivel NACIONAL

Frecuencia del Factor de Expansión
a nivel NACIONAL (Continuación)

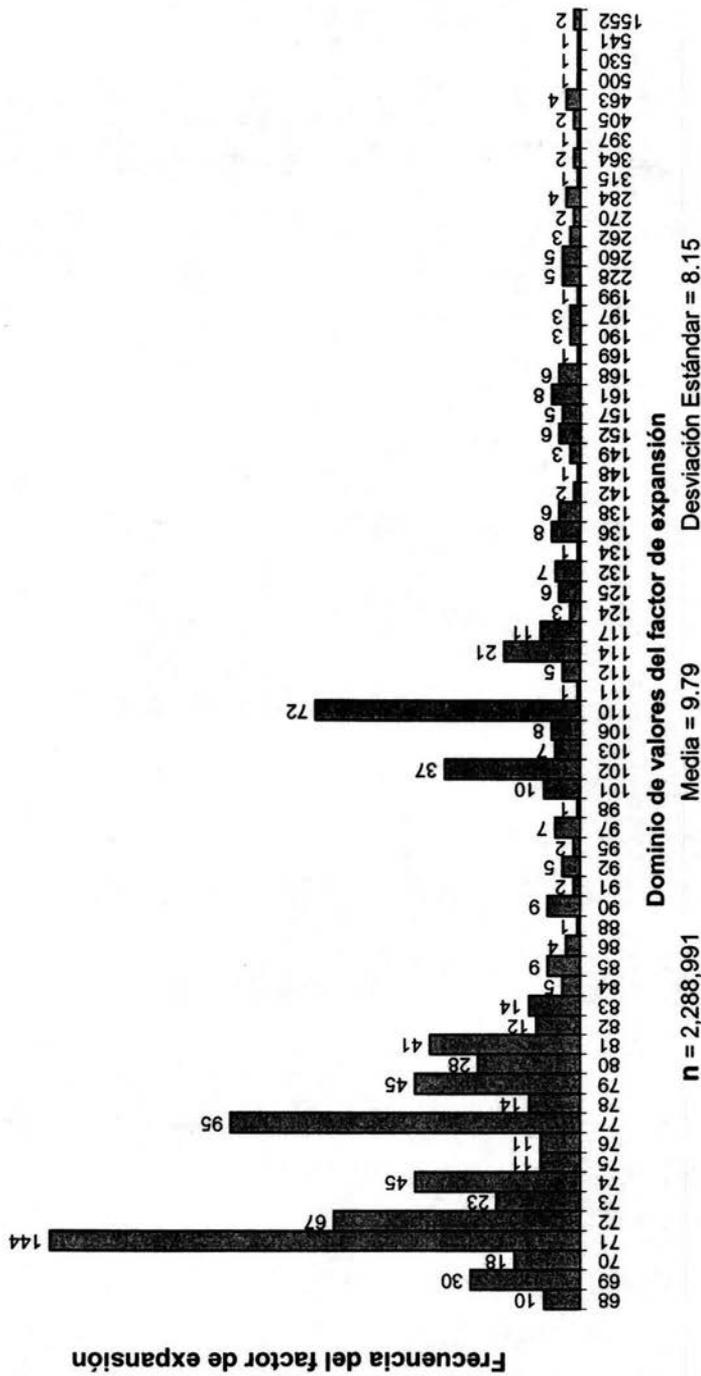


Figura A2.33 Frecuencia de los valores tomados por el Factor de expansión a nivel NACIONAL (Continuación)

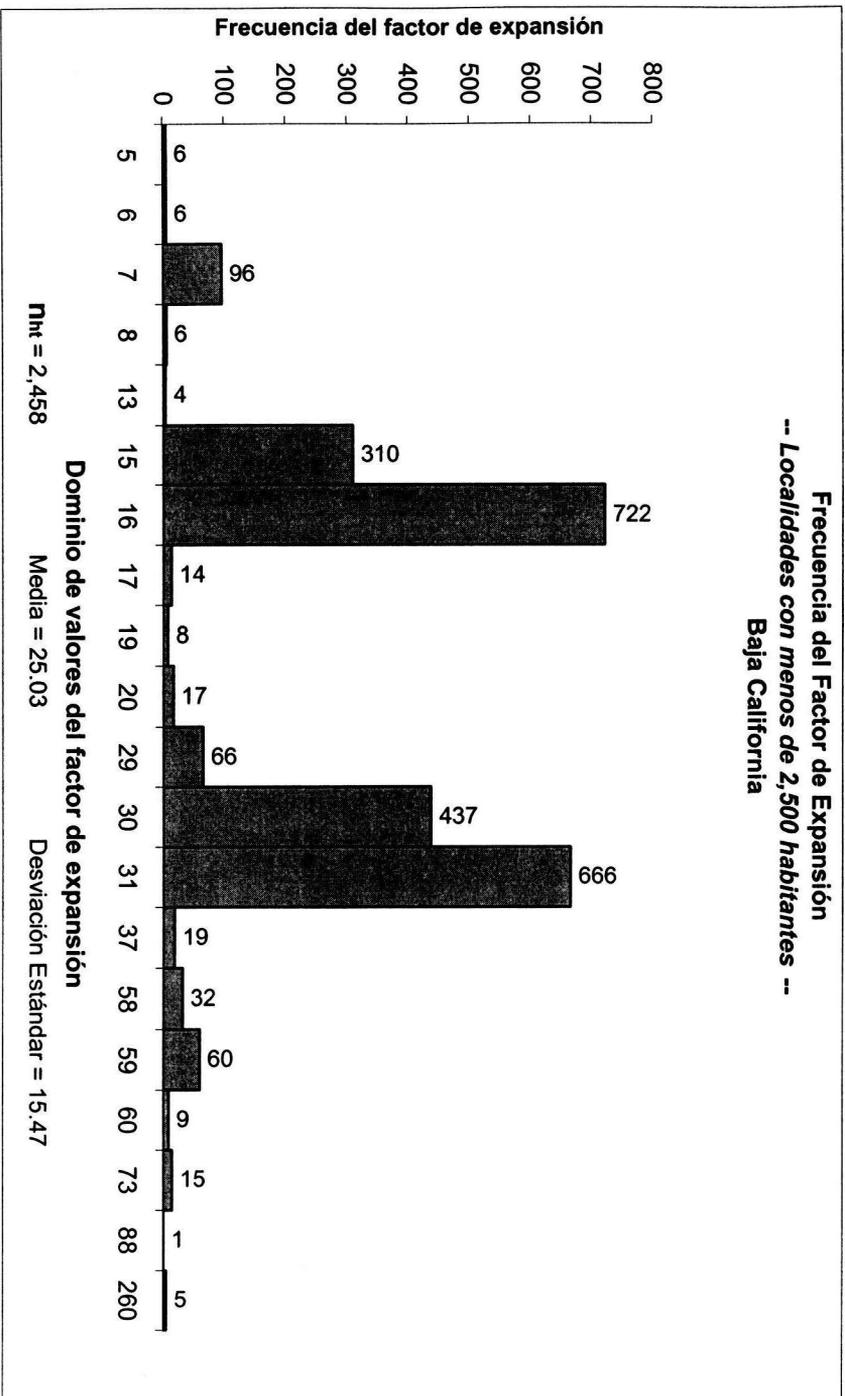


Figura A2.34 Frecuencia de los valores tomados por el Factor de expansión en Localidades con menos de 2,500 habitantes en Baja California

Frecuencia del Factor de Expansión
-- Localidades de 2,500 a 49,999 habitantes --
Baja California

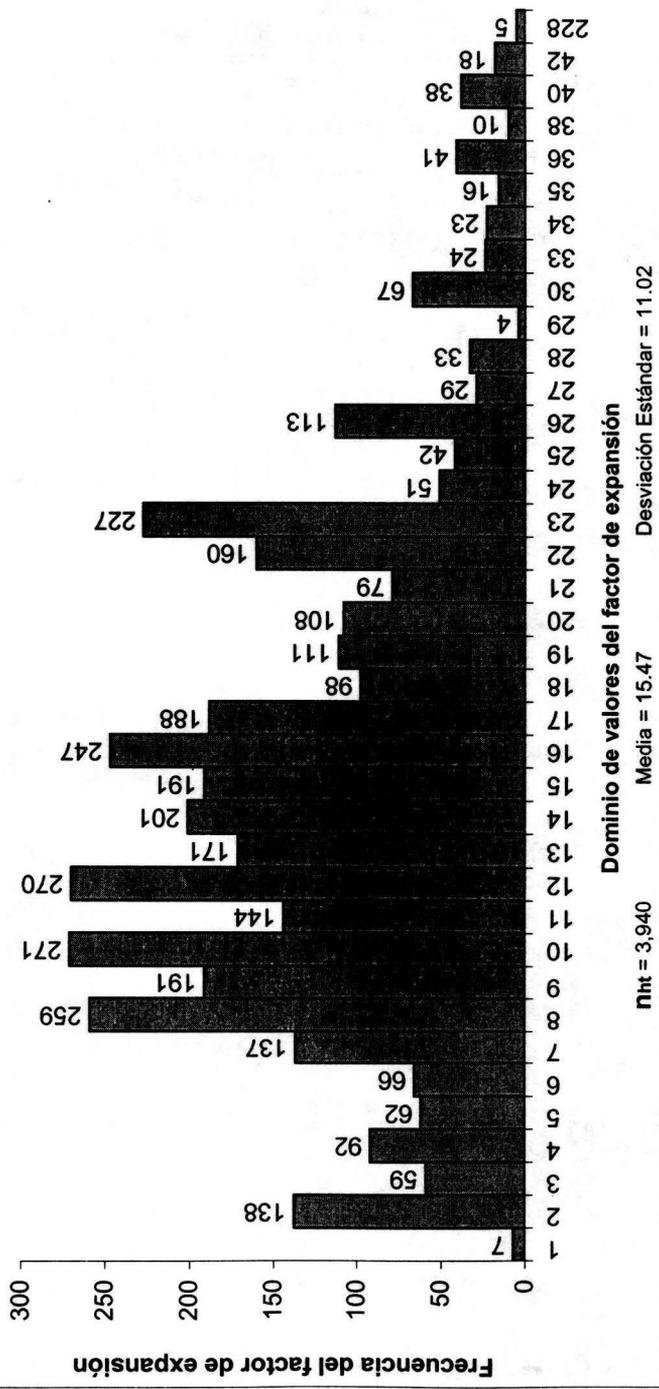


Figura A2.35 Frecuencia de los valores tomados por el Factor de expansión en Localidades de 2,500 a 49,999 habitantes en Baja California

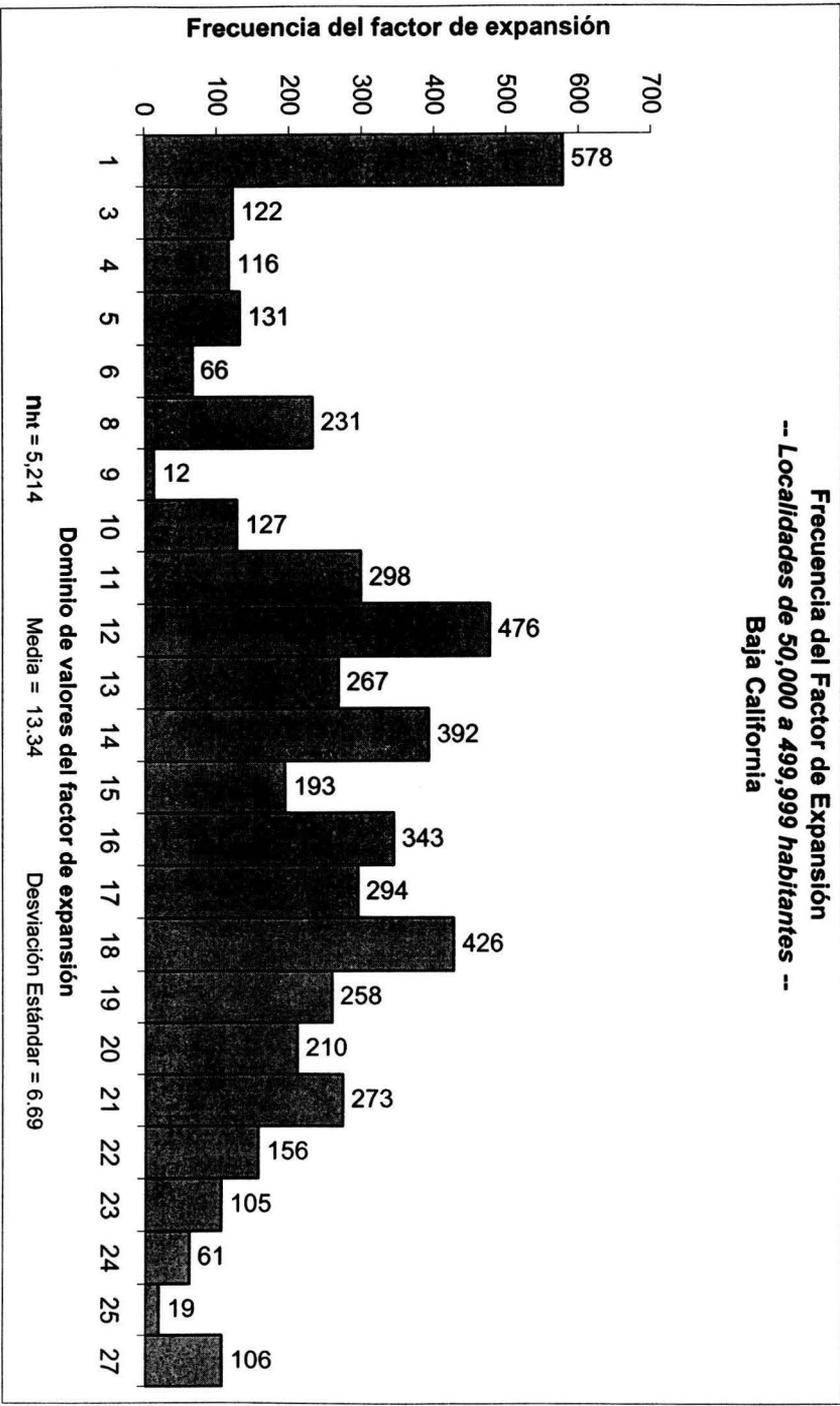


Figura A2.36 Frecuencia de los valores tomados por el Factor de expansión en Localidades de 50,000 a 499,999 habitantes en Baja California

Frecuencia del Factor de Expansión
 -- Localidades de 500,000 y más habitantes --
 Baja California

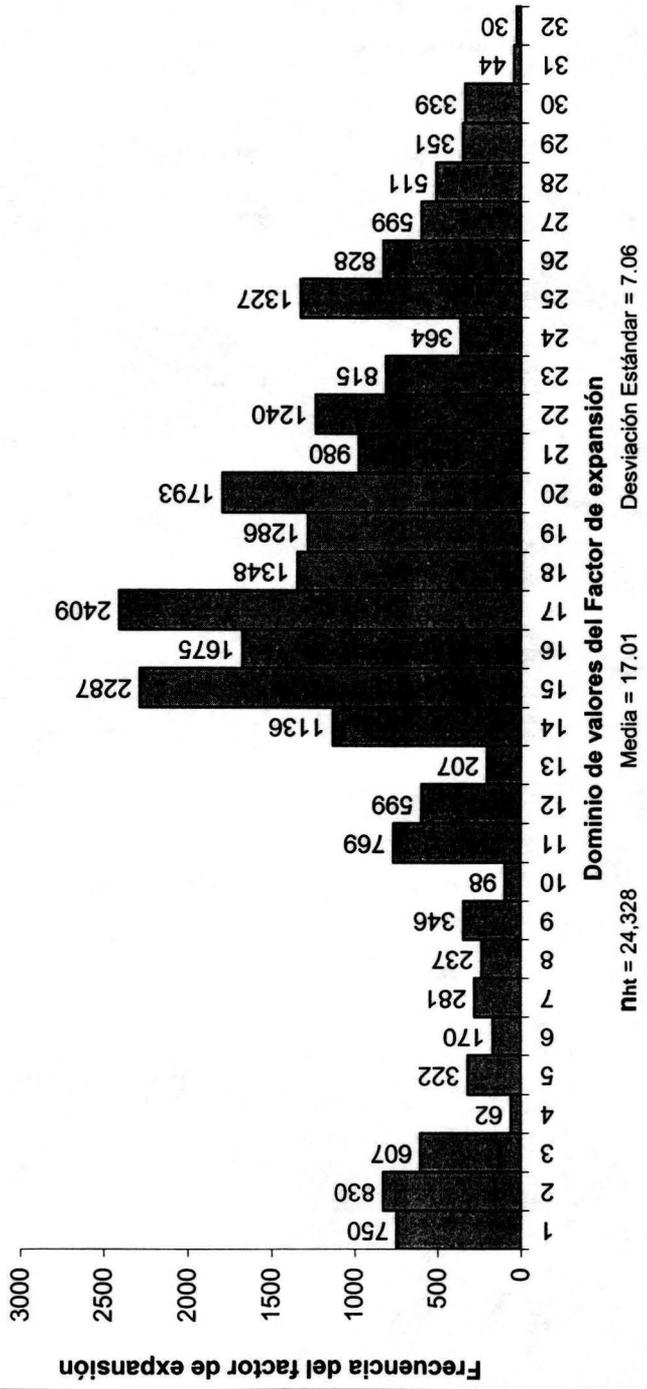


Figura A2.37 Frecuencia de los valores tomados por el Factor de expansión en Localidades de 500,000 y más habitantes en Baja California

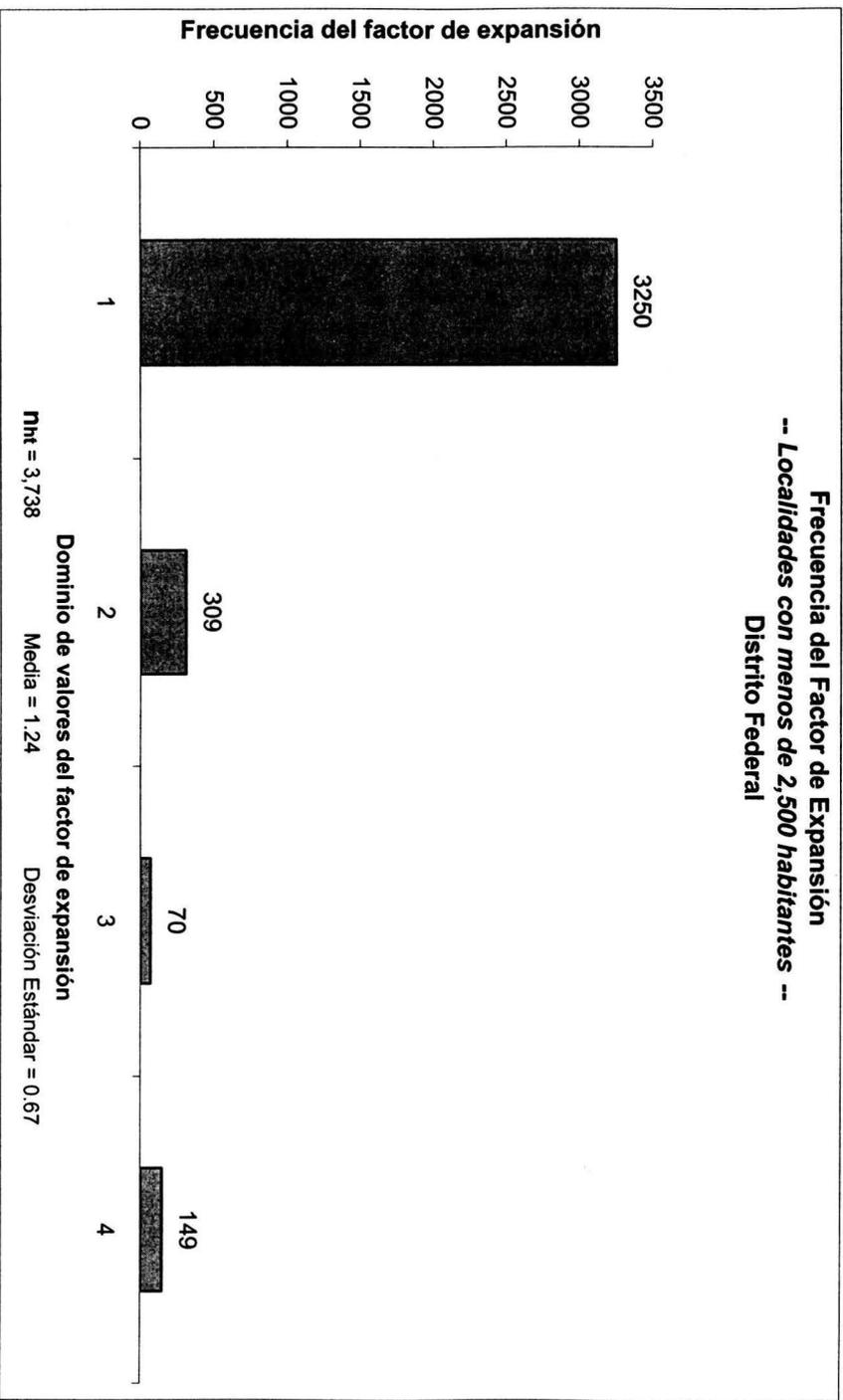


Figura A2.38 Frecuencia de los valores tomados por el Factor de expansión en Localidades con menos de 2,500 habitantes en el Distrito Federal

Frecuencia del Factor de Expansión
 -- Localidades de 2,500 a 49,999 habitantes --
 Distrito Federal

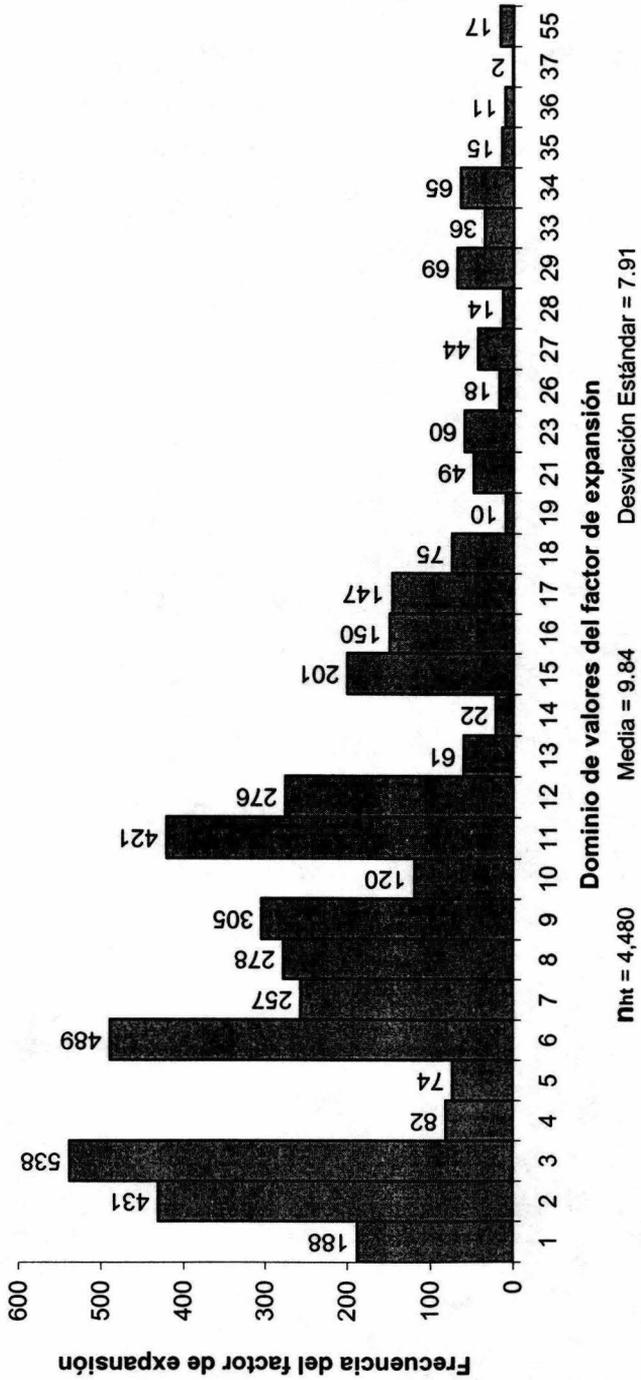


Figura A2.39 Frecuencia de los valores tomados por el Factor de expansión en Localidades de 2,500 a 49,999 habitantes en el Distrito Federal

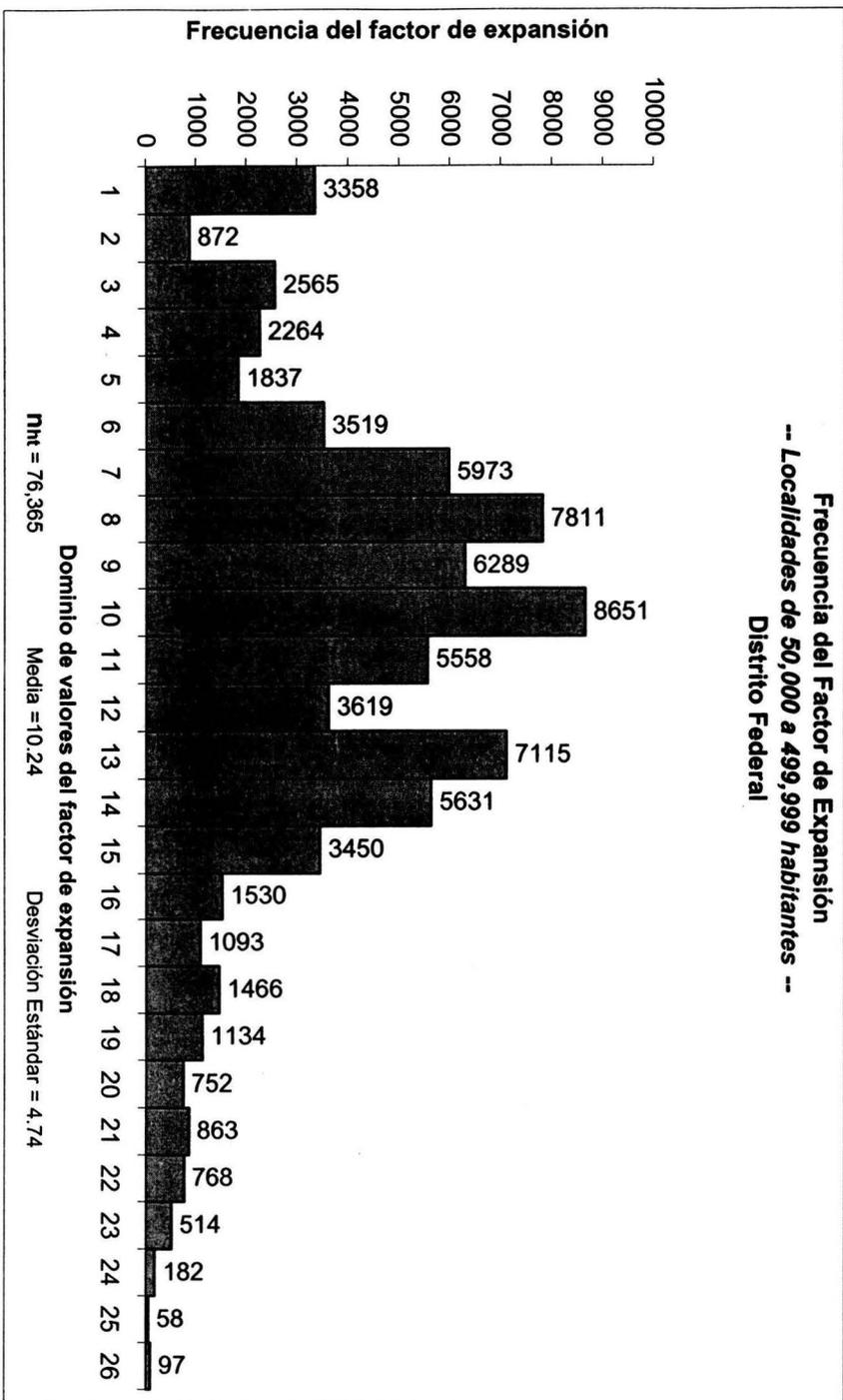


Figura A2.40 Frecuencia de los valores tomados por el Factor de expansión en Localidades de 50,000 a 499,999 habitantes en el Distrito Federal

Frecuencia del Factor de Expansión
-- Localidades de 500,000 y más habitantes --
Distrito Federal

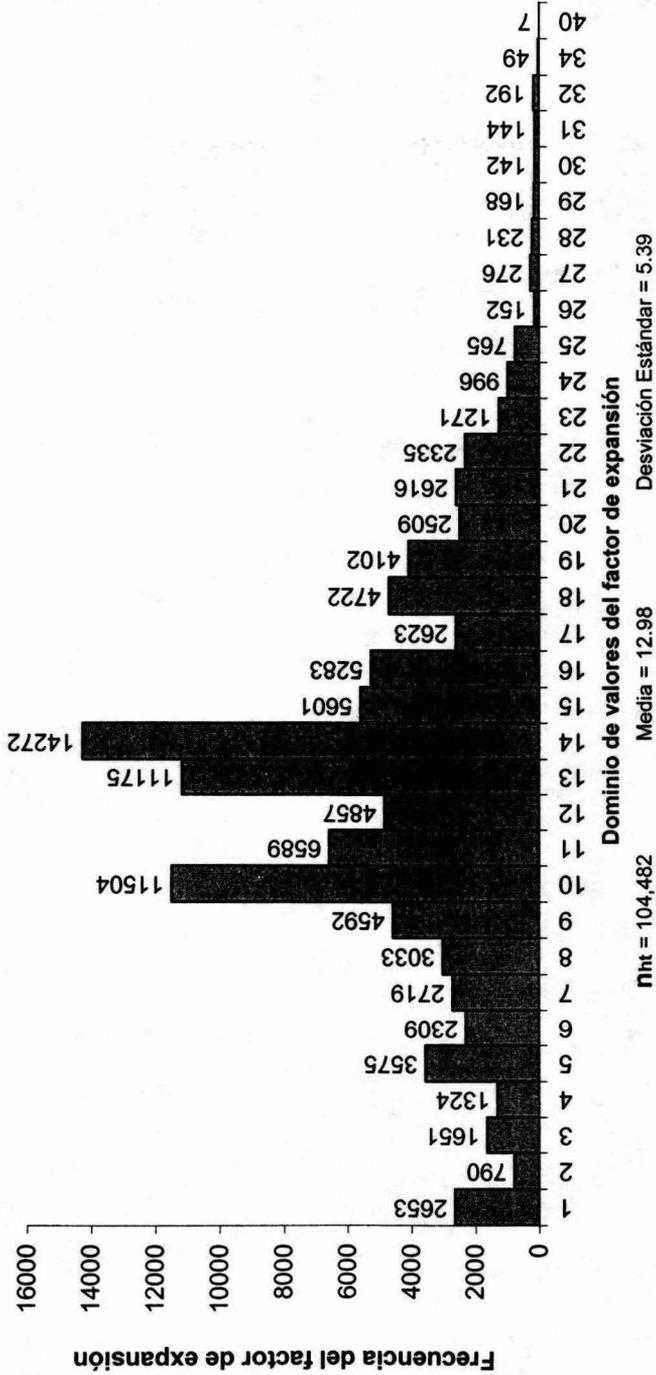


Figura A2.41 Frecuencia de los valores tomados por el Factor de expansión en Localidades de 500,000 y más habitantes en el Distrito Federal

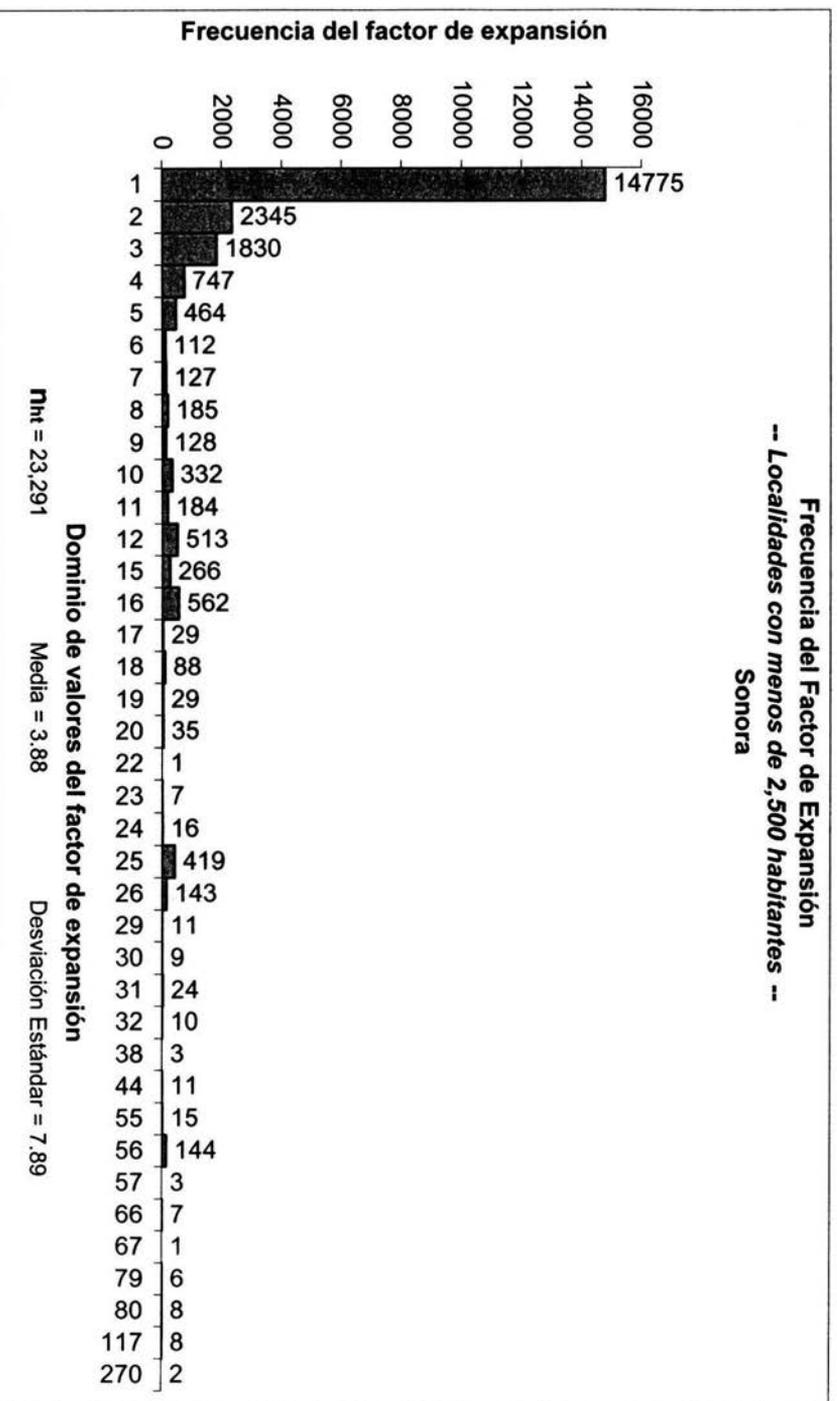


Figura A2.42 Frecuencia de los valores tomados por el Factor de expansión en Localidades con menos de 2,500 habitantes en Sonora

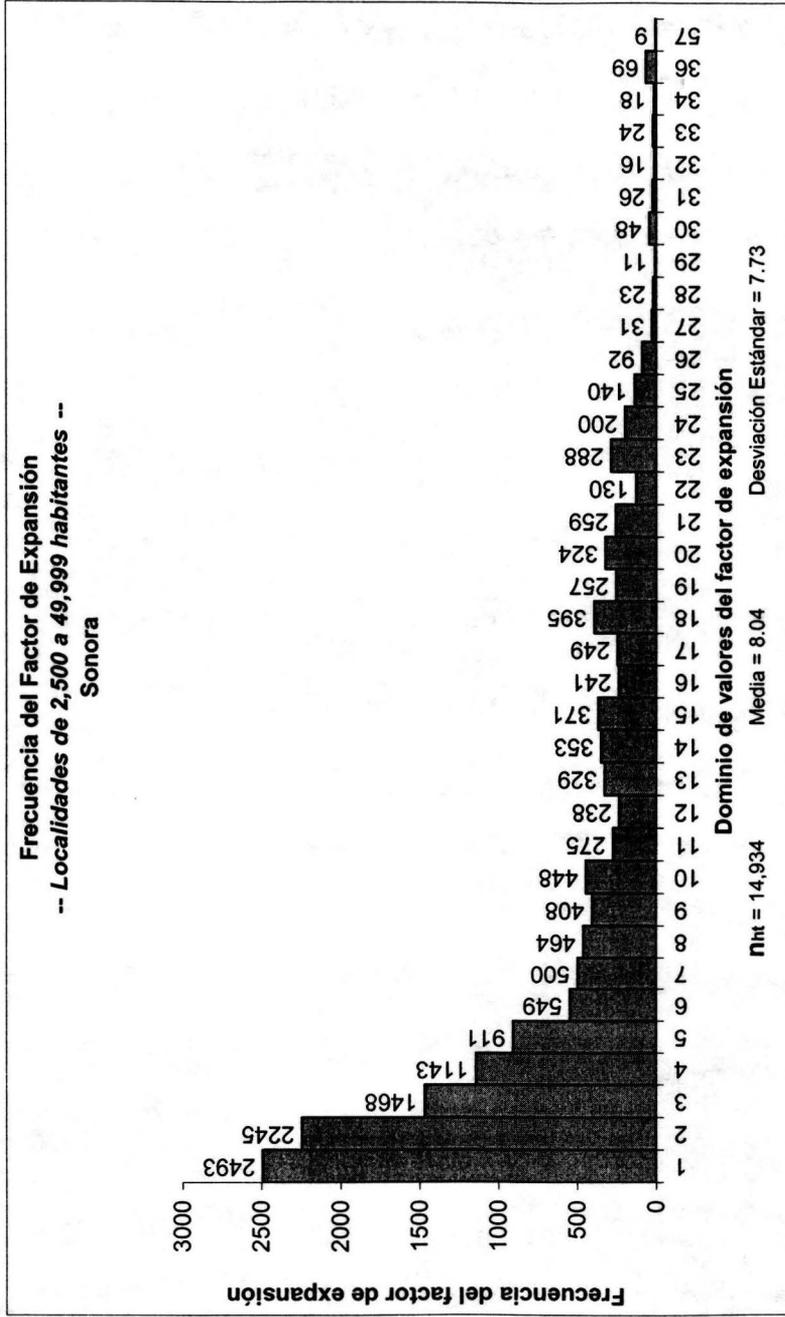


Figura A2.43 Frecuencia de los valores tomados por el Factor de expansión en Localidades de 2,500 a 49,999 habitantes en Sonora

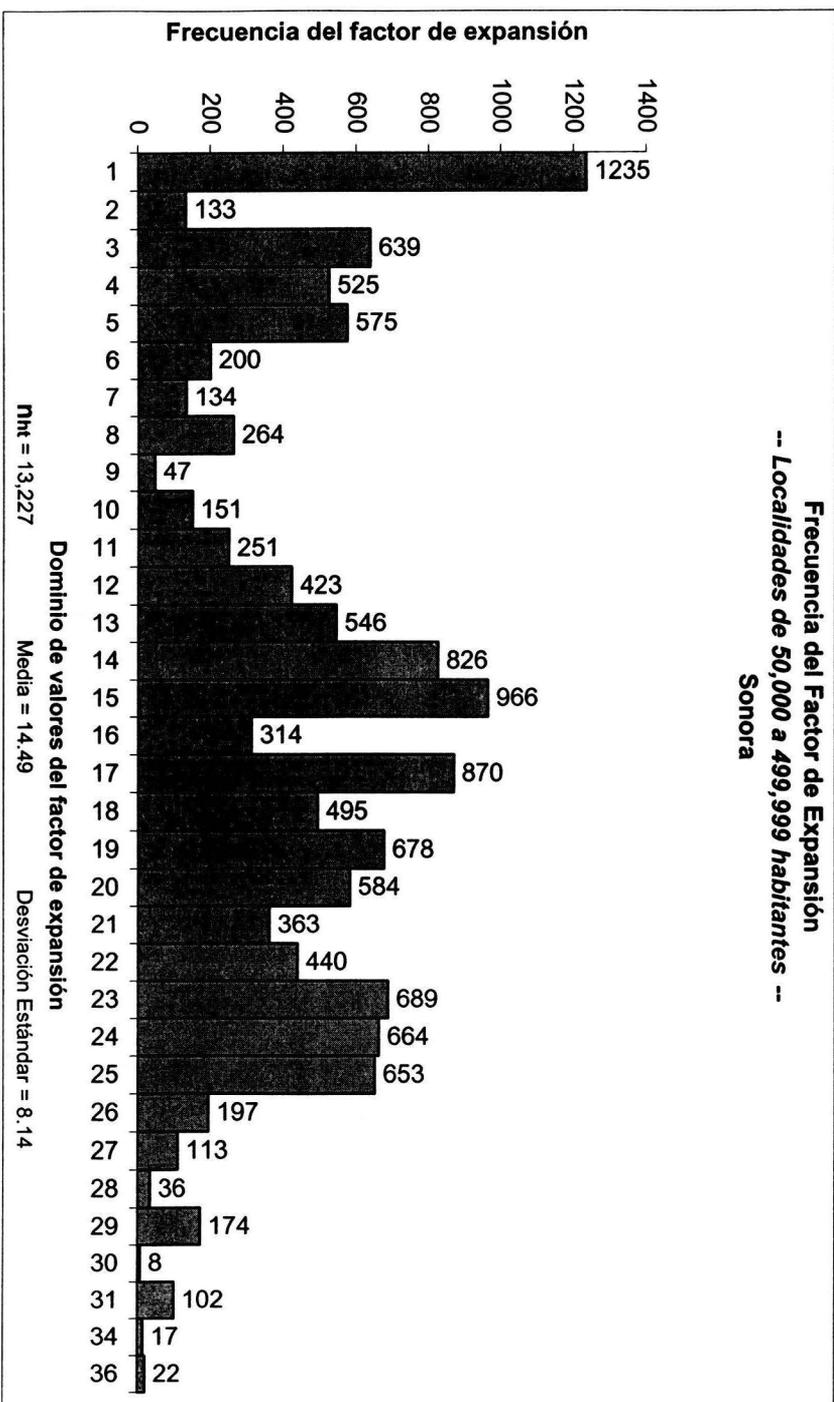


Figura A2.44 Frecuencia de los valores tomados por el Factor de expansión en Localidades de 50,000 a 499,999 habitantes en Sonora

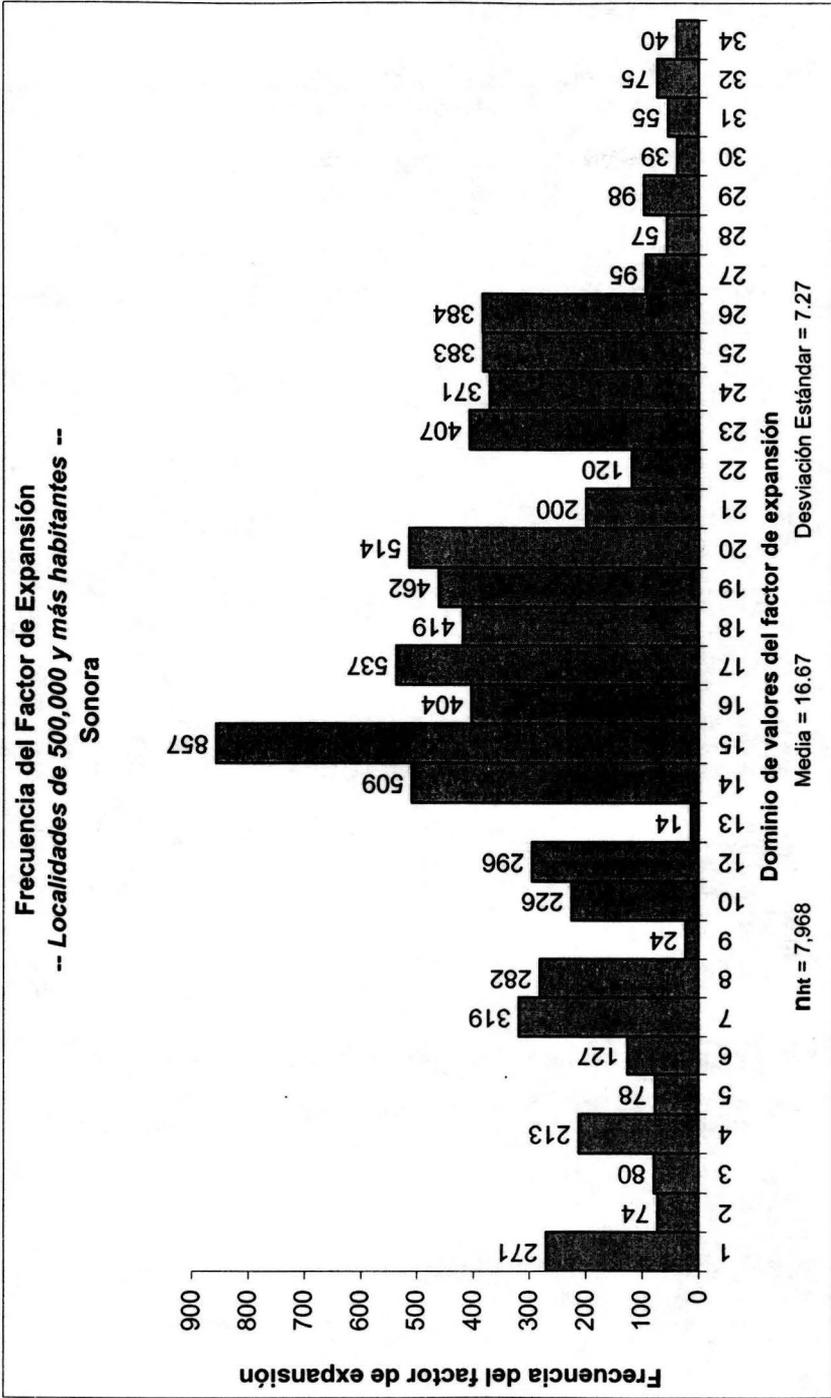


Figura A2.45 Frecuencia de los valores tomados por el Factor de expansión en Localidades de 500,000 y más habitantes en Sonora

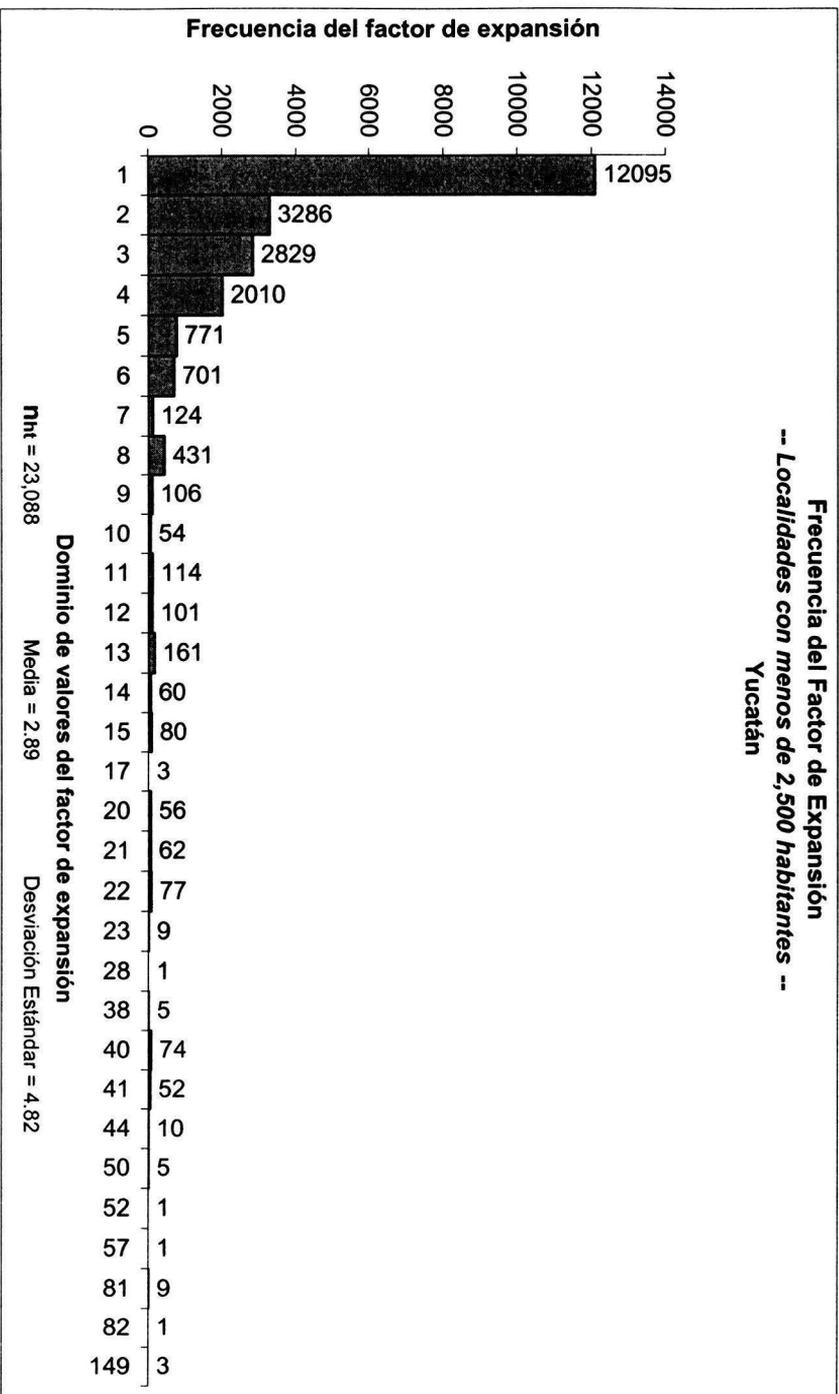


Figura A2.46 Frecuencia de los valores tomados por el Factor de expansión en Localidades con menos de 2,500 habitantes en Yucatán

Frecuencia del Factor de Expansión
 -- Localidades de 2,500 a 49,999 habitantes --
 Yucatán

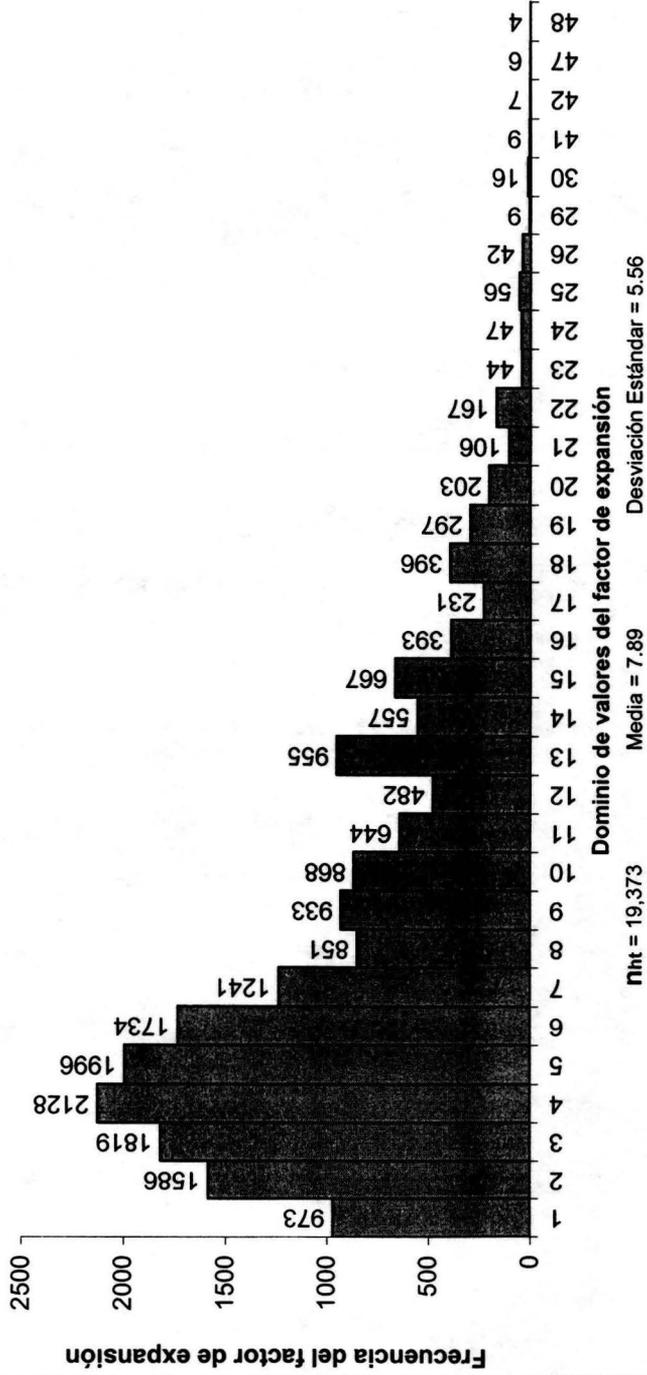


Figura A2.47 Frecuencia de los valores tomados por el Factor de expansión en Localidades de 2,500 a 49,999 habitantes en Yucatán

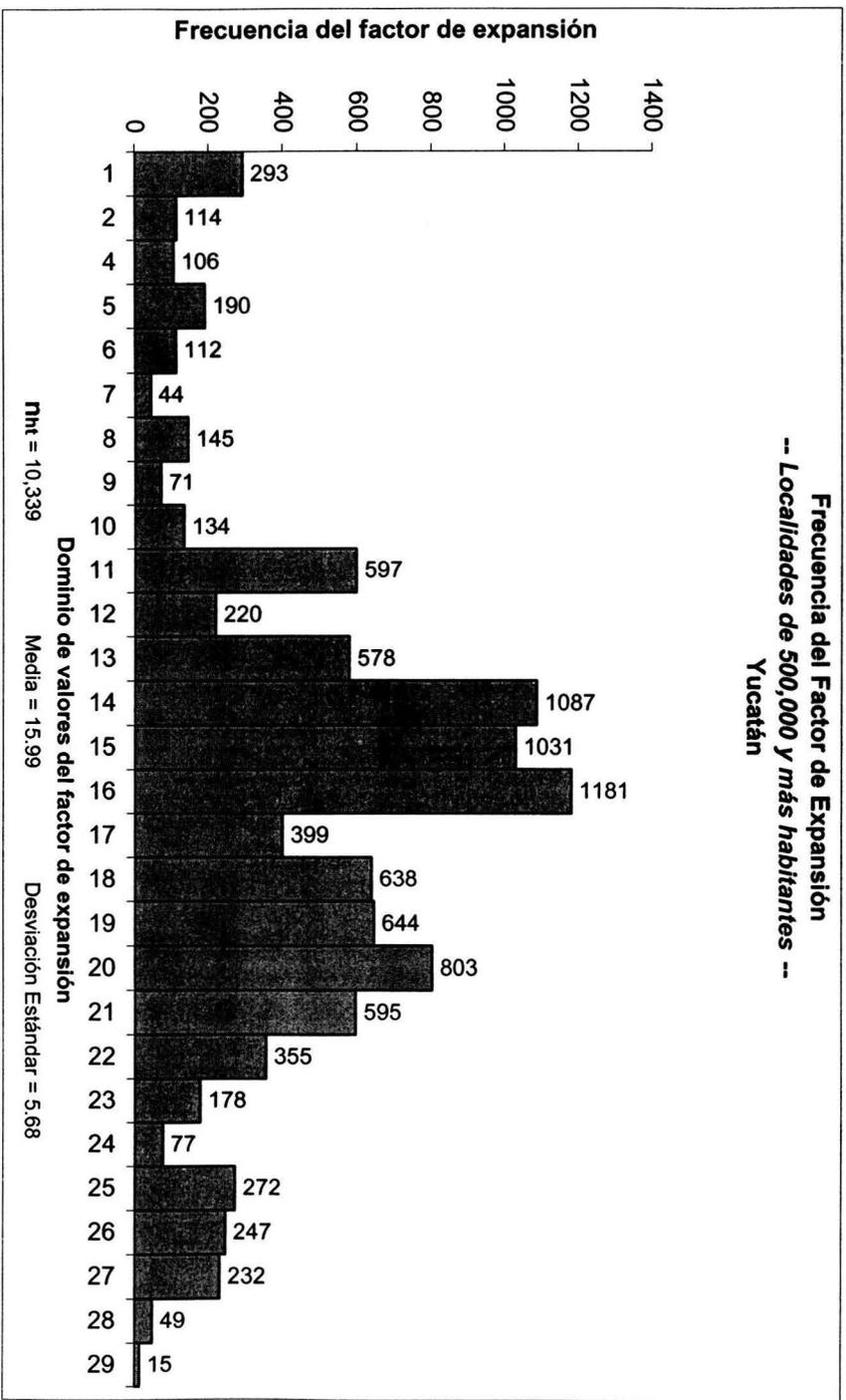


Figura A2.48 Frecuencia de los valores tomados por el Factor de expansión en Localidades de 500,000 y más habitantes en Yucatán

Anexo 3

Código utilizado en Sudaan

Para utilizar Sudaan y obtener los cálculos que se requieren cuando se considera el diseño muestral, es necesario que cada uno de los estratos de las bases de datos con las que se trabaja, cuenten con al menos dos UPM con clave distinta. Lo anterior se soluciona utilizando el programa que se presenta en el anexo 5.

Código para el cálculo del estimador de razón y su respectiva varianza (Series de Taylor) considerando el diseño de la muestra por entidad federativa

```
PROC RATIO DATA = "c:\\temp\\Ags\\vho_ags.sav"  
  FILETYPE=SPSS  DESIGN=WR;  
NEST ESTRATO UPM;  
WEIGHT FACTOR;  
NUMER TEL;  
DENOM DENOTEL;  
SETENV COLWIDTH=20;  
SETENV DECWIDTH=6;
```

Código para el cálculo del estimador de razón y su respectiva varianza (Jackknife) considerando el diseño de la muestra por entidad federativa

```
PROC RATIO DATA = "c:\\temp\\Ags\\vho_ags.sav"  
  FILETYPE=SPSS  DESIGN=JACKKNIFE;  
NEST ESTRATO UPM;
```

```
WEIGHT FACTOR;  
NUMER TEL;  
DENOM DENOTEL;  
SETENV COLWIDTH=20;  
SETENV DECWIDTH=6;
```

Código para el cálculo del estimador de razón y su respectiva varianza ignorando el diseño de muestra (m.a.s.) por entidad federativa

```
PROC RATIO DATA = "c:\\temp\\Ags\\vho_ag.sav"  
  FILETYPE=SPSS  DESIGN=SRS;  
NUMER TEL;  
DENOM DENOTEL;  
SETENV COLWIDTH=20;  
SETENV DECWIDTH=6;
```

Código para el cálculo del estimador de razón y su respectiva varianza (Series de Taylor) considerando el diseño de la muestra por rango de tamaños de localidades

```
PROC RATIO DATA = "c:\\temp\\Ags\\vho_ag1.sav"  
  FILETYPE=SPSS  DESIGN=WR;  
NEST ESTRATO UPM;  
WEIGHT FACTOR;  
NUMER TEL;  
DENOM DENOTEL;  
SETENV COLWIDTH=20;  
SETENV DECWIDTH=6;
```

Código para el cálculo del estimador de razón y su respectiva varianza ignorando el diseño de la muestra (m.a.s.) por rango de tamaños de localidades

```
PROC RATIO DATA = "c:\\temp\\Ags\\vho_ag1.sav"  
  FILETYPE=SPSS  DESIGN=SRS;  
NUMER TEL;  
DENOM DENOTEL;  
SETENV COLWIDTH=20;  
SETENV DECWIDTH=6;
```

Distribución del número de hogares por vivienda en Aguascalientes

		Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Número de hogares por vivienda	1	17696	96.6	96.6	96.6
	2	489	2.7	2.7	99.2
	3	111	.6	.6	99.8
	4	25	.1	.1	100.0
	5	5	.0	.0	100.0
	Total	18326	100.0	100.0	

Distribución del número de hogares por vivienda en Baja California

		Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Número de hogares por vivienda	1	35166	98.4	98.4	98.4
	2	470	1.3	1.3	99.8
	3	59	.2	.2	99.9
	4	24	.1	.1	100.0
	5	1	.0	.0	100.0
	6	1	.0	.0	100.0
	Total	35721	100.0	100.0	

Distribución del número de hogares por vivienda en Baja California Sur

		Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Número de hogares por vivienda	1	8917	98.1	98.1	98.1
	2	144	1.6	1.6	99.6
	3	24	.3	.3	99.9
	4	6	.1	.1	100.0
	5	2	.0	.0	100.0
	Total	9093	100.0	100.0	

Distribución del número de hogares por vivienda en Campeche

		Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Número de hogares por vivienda	1	14748	97.2	97.2	97.2
	2	370	2.4	2.4	99.6
	3	52	.3	.3	99.9
	4	8	.1	.1	100.0
	5	2	.0	.0	100.0
	Total	15180	100.0	100.0	

Distribución del número de hogares por vivienda en Coahuila

		Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Número de hogares por vivienda	1	45696	98.3	98.3	98.3
	2	680	1.5	1.5	99.7
	3	96	.2	.2	100.0
	4	18	.0	.0	100.0
	5	1	.0	.0	100.0
	6	2	.0	.0	100.0
	Total	46493	100.0	100.0	

Distribución del número de hogares por vivienda en Colima

		Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Número de hogares por vivienda	1	12260	97.9	97.9	97.9
	2	233	1.9	1.9	99.7
	3	26	.2	.2	99.9
	4	5	.0	.0	100.0
	5	1	.0	.0	100.0
	6	1	.0	.0	100.0
	Total	12526	100.0	100.0	

Distribución del número de hogares por vivienda en Chiapas

		Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Número de hogares por vivienda	1	90161	96.9	96.9	96.9
	2	2496	2.7	2.7	99.5
	3	359	.4	.4	99.9
	4	53	.1	.1	100.0
	5	12	.0	.0	100.0
	6	5	.0	.0	100.0
	7	1	.0	.0	100.0
	Total	93087	100.0	100.0	

Distribución del número de hogares por vivienda en Chihuahua

		Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Número de hogares por vivienda	1	71277	98.9	98.9	98.9
	2	671	.9	.9	99.9
	3	73	.1	.1	100.0
	4	13	.0	.0	100.0
	5	2	.0	.0	100.0
	6	1	.0	.0	100.0
	Total	72037	100.0	100.0	

Distribución del número de hogares por vivienda en el Distrito Federal

		Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Número de hogares por vivienda	1	178854	97.3	97.3	97.3
	2	3838	2.1	2.1	99.4
	3	877	.5	.5	99.8
	4	213	.1	.1	100.0
	5	58	.0	.0	100.0
	6	19	.0	.0	100.0
	7	8	.0	.0	100.0
	8	1	.0	.0	100.0
	9	2	.0	.0	100.0
	Total	183870	100.0	100.0	

Distribución del número de hogares por vivienda en Durango

		Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Número de hogares por vivienda	1	34523	98.4	98.4	98.4
	2	486	1.4	1.4	99.8
	3	60	.2	.2	100.0
	4	9	.0	.0	100.0
	5	1	.0	.0	100.0
	Total	35079	100.0	100.0	

Distribución del número de hogares por vivienda en Guanajuato

		Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Número de hogares por vivienda	1	72232	94.9	94.9	94.9
	2	2881	3.8	3.8	98.7
	3	752	1.0	1.0	99.7
	4	196	.3	.3	99.9
	5	43	.1	.1	100.0
	6	18	.0	.0	100.0
	7	3	.0	.0	100.0
	8	3	.0	.0	100.0
Total	76128	100.0	100.0		

Distribución del número de hogares por vivienda en Guerrero

		Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Número de hogares por vivienda	1	71405	97.2	97.2	97.2
	2	1728	2.4	2.4	99.5
	3	272	.4	.4	99.9
	4	49	.1	.1	100.0
	5	9	.0	.0	100.0
	6	3	.0	.0	100.0
	7	1	.0	.0	100.0
Total	73467	100.0	100.0		

Distribución del número de hogares por vivienda en Hidalgo

		Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Número de hogares por vivienda	1	59372	98.0	98.0	98.0
	2	1052	1.7	1.7	99.7
	3	132	.2	.2	99.9
	4	24	.0	.0	100.0
	5	4	.0	.0	100.0
	6	2	.0	.0	100.0
	7	1	.0	.0	100.0
	9	1	.0	.0	100.0
	Total	60588	100.0	100.0	

Distribución del número de hogares por vivienda en Jalisco

		Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Número de hogares por vivienda	1	139314	96.9	96.9	96.9
	2	3547	2.5	2.5	99.4
	3	678	.5	.5	99.9
	4	122	.1	.1	100.0
	5	24	.0	.0	100.0
	6	9	.0	.0	100.0
	7	2	.0	.0	100.0
	8	1	.0	.0	100.0
	9	1	.0	.0	100.0
Total	143698	100.0	100.0		

Distribución del número de hogares por vivienda en México

		Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Número de hogares por vivienda	1	236025	97.0	97.0	97.0
	2	5771	2.4	2.4	99.3
	3	1242	.5	.5	99.8
	4	300	.1	.1	100.0
	5	76	.0	.0	100.0
	6	10	.0	.0	100.0
	7	4	.0	.0	100.0
	Total	243428	100.0	100.0	

Distribución del número de hogares por vivienda en Michoacán

		Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Número de hogares por vivienda	1	97764	96.6	96.6	96.6
	2	2725	2.7	2.7	99.3
	3	557	.6	.6	99.9
	4	90	.1	.1	100.0
	5	25	.0	.0	100.0
	6	9	.0	.0	100.0
	7	3	.0	.0	100.0
	8	1	.0	.0	100.0
	Total	101174	100.0	100.0	

Distribución del número de hogares por vivienda en Morelos

		Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Número de hogares por vivienda	1	35249	97.4	97.4	97.4
	2	825	2.3	2.3	99.6
	3	111	.3	.3	99.9
	4	19	.1	.1	100.0
	5	1	.0	.0	100.0
	6	1	.0	.0	100.0
	Total	36206	100.0	100.0	

Distribución del número de hogares por vivienda en Nayarit

		Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Número de hogares por vivienda	1	18447	98.8	98.8	98.8
	2	190	1.0	1.0	99.8
	3	23	.1	.1	99.9
	4	9	.0	.0	100.0
	5	1	.0	.0	100.0
	Total	18670	100.0	100.0	

Distribución del número de hogares por vivienda en Nuevo León

		Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Número de hogares por vivienda	1	77415	97.2	97.2	97.2
	2	1783	2.2	2.2	99.5
	3	323	.4	.4	99.9
	4	69	.1	.1	100.0
	5	18	.0	.0	100.0
	6	5	.0	.0	100.0
	Total	79613	100.0	100.0	

Distribución del número de hogares por vivienda en Oaxaca

		Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Número de hogares por vivienda	1	150332	97.7	97.7	97.7
	2	3072	2.0	2.0	99.7
	3	416	.3	.3	99.9
	4	75	.0	.0	100.0
	5	16	.0	.0	100.0
	6	2	.0	.0	100.0
	7	1	.0	.0	100.0
Total	153914	100.0	100.0		

Distribución del número de hogares por vivienda en Puebla

		Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Número de hogares por vivienda	1	124673	97.0	97.0	97.0
	2	3218	2.5	2.5	99.5
	3	532	.4	.4	99.9
	4	117	.1	.1	100.0
	5	28	.0	.0	100.0
	6	4	.0	.0	100.0
	7	3	.0	.0	100.0
Total	128575	100.0	100.0		

Distribución del número de hogares por vivienda en Querétaro

		Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Número de hogares por vivienda	1	26047	96.6	96.6	96.6
	2	686	2.5	2.5	99.1
	3	172	.6	.6	99.8
	4	42	.2	.2	99.9
	5	13	.0	.0	100.0
	6	6	.0	.0	100.0
	7	1	.0	.0	100.0
	Total	26967	100.0	100.0	

Distribución del número de hogares por vivienda en Quintana Roo

		Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Número de hogares por vivienda	1	15653	97.8	97.8	97.8
	2	275	1.7	1.7	99.6
	3	53	.3	.3	99.9
	4	13	.1	.1	100.0
	5	4	.0	.0	100.0
	6	1	.0	.0	100.0
	Total	15999	100.0	100.0	

Distribución del número de hogares por vivienda en San Luis Potosí

		Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Número de hogares por vivienda	1	55358	97.7	97.7	97.7
	2	1062	1.9	1.9	99.6
	3	193	.3	.3	99.9
	4	42	.1	.1	100.0
	5	11	.0	.0	100.0
	6	2	.0	.0	100.0
	Total	56668	100.0	100.0	

Distribución del número de hogares por vivienda en Sinaloa

		Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Número de hogares por vivienda	1	40304	98.4	98.4	98.4
	2	587	1.4	1.4	99.8
	3	65	.2	.2	100.0
	4	11	.0	.0	100.0
	5	7	.0	.0	100.0
	7	1	.0	.0	100.0
	Total	40975	100.0	100.0	

Distribución del número de hogares por vivienda en Sonora

		Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Número de hogares por vivienda	1	58545	98.8	98.8	98.8
	2	609	1.0	1.0	99.9
	3	58	.1	.1	100.0
	4	15	.0	.0	100.0
	5	1	.0	.0	100.0
	6	1	.0	.0	100.0
	Total	59229	100.0	100.0	

Distribución del número de hogares por vivienda en Tabasco

		Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Número de hogares por vivienda	1	37579	97.1	97.1	97.1
	2	957	2.5	2.5	99.6
	3	141	.4	.4	99.9
	4	24	.1	.1	100.0
	5	4	.0	.0	100.0
	6	1	.0	.0	100.0
	Total	38706	100.0	100.0	

Distribución del número de hogares por vivienda en Tamaulipas

		Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Número de hogares por vivienda	1	56294	98.6	98.6	98.6
	2	708	1.2	1.2	99.8
	3	96	.2	.2	100.0
	4	14	.0	.0	100.0
	5	9	.0	.0	100.0
	9	1	.0	.0	100.0
	Total	57122	100.0	100.0	

Distribución del número de hogares por vivienda en Tlaxcala

		Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Número de hogares por vivienda	1	27866	95.7	95.7	95.7
	2	1015	3.5	3.5	99.2
	3	186	.6	.6	99.8
	4	35	.1	.1	100.0
	5	9	.0	.0	100.0
	Total	29111	100.0	100.0	

Distribución del número de hogares por vivienda en Veracruz

		Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Número de hogares por vivienda	1	178761	97.7	97.7	97.7
	2	3688	2.0	2.0	99.7
	3	490	.3	.3	100.0
	4	66	.0	.0	100.0
	5	11	.0	.0	100.0
	6	5	.0	.0	100.0
	9	2	.0	.0	100.0
		Total	183023	100.0	100.0

Distribución del número de hogares por vivienda en Yucatán

		Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Número de hogares por vivienda	1	48563	95.9	95.9	95.9
	2	1748	3.5	3.5	99.3
	3	298	.6	.6	99.9
	4	44	.1	.1	100.0
	5	6	.0	.0	100.0
	7	1	.0	.0	100.0
	Total	50660	100.0	100.0	

Distribución del número de hogares por vivienda en Zacatecas

		Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Número de hogares por vivienda	1	41878	97.8	97.8	97.8
	2	784	1.8	1.8	99.6
	3	129	.3	.3	99.9
	4	24	.1	.1	100.0
	5	3	.0	.0	100.0
	6	2	.0	.0	100.0
	Total	42820	100.0	100.0	

Distribución del número de hogares por vivienda a nivel NACIONAL

		Frecuencia	Porcentaje	Porcentaje Válido	Porcentaje Acumulado
Número de hogares por vivienda	1	2178374	97.33	97.33	97.33
	2	48788	2.18	2.18	99.51
	3	8656	0.39	0.39	99.90
	4	1774	0.08	0.08	99.97
	5	408	0.02	0.02	99.99
	6	115	0.01	0.01	100.00
	7	30	0.00	0.00	100.00
	8	6	0.00	0.00	100.00
	9	7	0.00	0.00	100.00
Total	2238158	100.00	100.00		

Anexo 5

Programa en C utilizado para preparar los archivos antes de utilizar Sudaan

Cuando se utiliza Sudaan y se emplea el diseño de la muestra por Series de Taylor o Jackknife, es necesario que se cuente con al menos dos UPM con clave distinta por cada estrato, de lo contrario Sudaan indica error y no ejecuta el programa. Debido a que en los archivos originales había varios estratos con una sola UPM se tuvo que realizar el siguiente programa escrito en lenguaje de programación 'C', el cual soluciona el problema cambiando las claves del estrato en donde hay una sola UPM por las claves del estrato anterior y las claves de la UPM que se encontraba como única toman la clave (el número) de la UPM más grande del estrato anterior +1. A nivel estatal, el rango del número de estratos con una sola UPM que fueron colapsados varía de 21 a más de 1500 correspondientes a los estados de Baja California Sur y México respectivamente.

```
#include <stdio.h>
#include <string.h>
#include <stdlib.h>
#include <conio.h>

void main(void)
{
    FILE *original, *generado, *bitacora;
```

```
char ArchivoLinea[21], *PosicionComa,
EstratoOriginal[20], UpmOriginal[20];
long Linea;
int nCaracteres;
int UPM, UpmMayor, UpmAnteriorMayor,
DiferentesUpmXEstrato, UpmRepetidos;
long ESTRATO, EstratoUltimo, EstratoAnterior,
InicioEstrato, FinEstrato;
int NumeroEstrato;
int ArregloUpm[100][2];
int Fila, Repetir;

clrscr();

original=fopen("c:\\prueba.csv", "rt");
generado=fopen("c:\\generado.txt", "wt");
bitacora=fopen("c:\\bitacora.txt", "wt");

if (original==NULL || generado==NULL || bitacora==NULL)
{
    clrscr();
    printf("\n\n Error al abrir los archivo \a");
    exit(0);
}

fgets(ArchivoLinea, 20, original);
Linea=1;
while(!feof(original))
{
    printf("%ld\n", Linea);
    PosicionComa=strchr(ArchivoLinea, ',');
    nCaracteres=PosicionComa-ArchivoLinea;

    strncpy(EstratoOriginal, ArchivoLinea,
nCaracteres);
    EstratoOriginal[nCaracteres]='\0';

    strncpy(UpmOriginal, PosicionComa+1,
strlen(ArchivoLinea)-nCaracteres-2);
    UpmOriginal[strlen(ArchivoLinea)-nCaracteres-
2]='\0';

    UPM=atoi(UpmOriginal);
    ESTRATO=atol(EstratoOriginal);
```

```

        if (Linea==1)
        {
            fprintf(generado, "%s,%s\n", EstratoOriginal,
UpmOriginal);
            fprintf(bitacora,"%-6s %-10s %-10s %-10s %-10s
%-5s %-10s %-8s\n", "Numero", "Estrato", "Inicio", "Fin",
"Diferentes", "Mayor", "Anterior", "Cambiado");
        }
        else
        {
            if (Linea==2)
            {
                EstratoUltimo=ESTRATO;
                EstratoAnterior=ESTRATO;
                InicioEstrato=Linea;
                FinEstrato=Linea;
                NumeroEstrato=1;
                UpmMayor=UPM;
                UpmAnteriorMayor=UPM;
                DiferentesUpmXEstrato=1;
                UpmRepetidos=1;
                ArregloUpm [DiferentesUpmXEstrato-
1] [0]=UPM;
                ArregloUpm [DiferentesUpmXEstrato-1] [1]=1;
            }
            else
            {
                if (ESTRATO==EstratoUltimo)
                {
                    FinEstrato=Linea;

                    if (UPM==ArregloUpm [DiferentesUpmXEstrato-1] [0])
                    {
                        UpmRepetidos++;

                        ArregloUpm [DiferentesUpmXEstrato-1] [1]=UpmRepetidos;
                    }
                    else
                    {
                        UpmRepetidos=1;
                        DiferentesUpmXEstrato++;

                        ArregloUpm [DiferentesUpmXEstrato-1] [0]=UPM;
                    }
                }
            }
        }
    }
}

```

```

ArregloUpm[DiferentesUpmXEstrato-1][1]=UpmRepetidos;

        if (UPM>UpmMayor)
        {
            UpmMayor=UPM;
        }
    }
else
{
    if(DiferentesUpmXEstrato>1)
    {
        for(Fila=0;
Fila<DiferentesUpmXEstrato; Fila++)
        {
            for(Repetir=1;
Repetir<=ArregloUpm[Fila][1]; Repetir++)
            {

                fprintf(generado,"%010ld,%04d\n", EstratoUltimo,
ArregloUpm[Fila][0]);

            }
        }
        fprintf(bitacora,"%-6d %010ld
%-10ld %-10ld %-10d %04d\n", NumeroEstrato, EstratoUltimo,
InicioEstrato, FinEstrato, DiferentesUpmXEstrato, UpmMayor);
        EstratoAnterior=EstratoUltimo;
        EstratoUltimo=ESTRATO;
        UpmAnteriorMayor=UpmMayor;
        UpmMayor=UPM;
    }
else
{
    for(Repetir=1;
Repetir<=UpmRepetidos; Repetir++)
    {

        fprintf(generado,"%010ld,%04d\n", EstratoAnterior,
UpmAnteriorMayor+1);

    }
    fprintf(bitacora,"%-6d %010ld
%-10ld %-10ld %-10d %04d %010ld %04d\n", NumeroEstrato,
EstratoUltimo, InicioEstrato, FinEstrato,

```

```

DiferentesUpmXEstrato, UpmMayor, EstratoAnterior,
UpmAnteriorMayor+1);

        EstratoUltimo=ESTRATO;
        UpmAnteriorMayor++;
        UpmMayor=UPM;
    }
    NumeroEstrato++;
    InicioEstrato=Linea;
    FinEstrato=Linea;
    DiferentesUpmXEstrato=1;
    UpmRepetidos=1;
    ArregloUpm[DiferentesUpmXEstrato-
1] [0]=UPM;
        ArregloUpm[DiferentesUpmXEstrato-
1] [1]=1;
    }
    }
    fgets(ArchivoLinea, 20, original);
    Linea++;
}

if(DiferentesUpmXEstrato>1)
{
    for(Fila=0; Fila<DiferentesUpmXEstrato; Fila++)
    {
        for(Repetir=1; Repetir<=ArregloUpm[Fila] [1];
Repetir++)
        {
            fprintf(generado,"%010ld,%04d\n",
EstratoUltimo, ArregloUpm[Fila] [0]);
        }
        fprintf(bitacora,"%-6d %010ld %-10ld %-10ld %-10d
%04d\n", NumeroEstrato, EstratoUltimo, InicioEstrato,
FinEstrato, DiferentesUpmXEstrato, UpmMayor);
    }
    else
    {
        for(Repetir=1; Repetir<=UpmRepetidos; Repetir++)
        {
            fprintf(generado,"%010ld,%04d\n",
EstratoAnterior, UpmAnteriorMayor+1);
        }
    }
}

```

```
        fprintf(bitacora,"%-6d %010ld %-10ld %-10ld %-10d
%04d %010ld %04d\n", NumeroEstrato, EstratoUltimo,
InicioEstrato, FinEstrato, DiferentesUpmXEstrato, UpmMayor,
EstratoAnterior, UpmAnteriorMayor+1);
    }
    fclose(original);
    fclose(generado);
    fclose(bitacora);
    printf("Concluido\n\a");
    printf("Oprime una tecla para terminar");
    getch();
}
```