UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

EVALUACIÓN DE LA COBERTURA DE GUARDERÍAS Y LA DEMANDA POTENCIAL DE NIÑOS A PARTIR DE TÉCNICAS DE MUESTREO PROBABILÍSTICO.

T E S I S

QUE PARA OBTENER EL TÍTULO DE

A C T U A R I A

P R E S E N T A

NORMA ALEJANDRA VERGARA LOPE GRACIA

Tutor:
ACT. JOSÉ FABIÁN GONZÁLEZ FLORES

2008

FACULTAD DE CIENCIAS
UNAM

1. Datos del alumno
    Vergara-Lope
    Gracia
    Norma Alejandra
    0445518542865
    Universidad Nacional Autónoma de México
    Facultad de Ciencias
    Actuaría
    099333599

2. Datos del tutor
    Actuario
    José Fabián
    González
    Flores

3. Datos del sinodal 1
    Maestra en Ciencias
    Jésica
    Hernández
    Rojano

4. Datos del sinodal 2
    Actuario
    Francisco
    Sánchez
    Villareal

5. Datos del sinodal 3
    Maestra en Demografía
    Patricia
    Fernández
    Ham

6. Datos del sinodal 4
    Actuario
    Jaime
    Vázquez
    Alamilla

7. Datos del trabajo escrito
    Evaluación de la Cobertura de Guarderías y la Demanda Potencial de Niños a Partir de Técnicas de Muestreo Probabilístico.
    85 páginas
    2008

AGRADECIMIENTOS

Expreso mi agradecimiento a la Universidad Nacional Autónoma de México por permitirme cristalizar mis proyectos, por la distinción con la cual me honran para compartir las experiencias que me ha brindado enormes satisfacciones,

A Dios, por ser mi guía en este camino, a El por estar siempre a mi lado

A mi Madre Que con su apoyo incondicional y amor logré ser lo que soy ahora, por creer en mí y motivarme en todo momento sin esperar nada a cambio y por ser la luz de mi vida

Gracias ¡Te Quiero Mucho!

A mi Padre Que con sus esfuerzos y cariño me dio la ayuda para salir adelante

A mi Hermano Por soportarme y estar junto a mi en cada momento.

Al Act. Fabián González Flores Por darme la confianza y el apoyo necesario para realizar este proyecto, dándome animo para seguir adelante

A mis Amigos Por su apoyo y cariño incondicional, por estar siempre en las buenas y en las malas, aunque algunos no estuvieron físicamente conmigo en todo momento pero moralmente siempre estuvieron ahí cuando yo los necesitaba, ellos saben quienes son.

A mis compañeros de trabajo que me dieron consejos y alientos para continuar y por hacer mi estancia allí una gran experiencia.

A todos los que no confiaron en mí, que gracias a ellos saque las fuerzas necesarias para seguir adelante y no quedarme en el camino.

INTRODUCCIÓN

El objetivo de este estudio es evaluar y analizar a nivel delegacional la oferta de los servicios de Guarderías y la demanda potencial de niños menores de 4 años, a partir de los elementos teóricos indispensables para el desarrollo de técnicas de muestreo probabilístico.

En la actualidad, uno de los desafíos que enfrentan las familias Mexicanas es el cuidado de los niños mientras los padres trabajan. A medida de que crezca la población laboral, también aumentará la demanda para servicios de guarderías, y aun en la actualidad estos servicios son insuficientes en Instituciones Públicas ya que no se tiene la capacidad requerida para otorgar esta prestación. Es por ello, que resulta útil construir un modelo de estimación de tasas de participación, con bases de técnicas de muestreo, que de como resultado un análisis estadístico sobre la situación de la demanda insatisfecha y así verificar si se cumplirán con las expectativas de los trabajadores que tienen derecho a la prestación, pues se espera que en un futuro, la cobertura sea ligeramente superior.

La Encuesta Nacional de Empleo y Seguridad Social 2004 (ENESS) mostró que sólo 32.8 por ciento de la población ocupada masculina y 35.3 por ciento de la femenina contaban con algún tipo de prestación laboral. Entre las prestaciones de la seguridad social están las guarderías; en éstas se brinda atención integral a los hijos de las trabajadoras y en casos especiales, de los trabajadores.

Aunque el número y capacidad de atención de estos servicios ha crecido de manera constante, aún son insuficientes, por lo que madres y padres trabajadores se han visto obligados a recurrir a las redes de apoyo familiar y social – como abuelos y vecinos –, a los servicios personales remunerados, o incluso, a dejar solos a sus hijos durante su jornada laboral.

Los resultados de la ENESS 2004 muestran que 16 por ciento de la población de seis años y menos fueron encargados al cuidado de alguna institución o persona distinta a sus padres y de éstos 9.3 por ciento fueron atendidos en guarderías públicas, 4.9 por ciento fueron atendidos en guarderías privadas, 56.8 por ciento fueron cuidados por un familiar, 5.6 por ciento fueron cuidados por una persona remunerada, 2.1 por ciento fueron cuidados por un no familiar no remunerado y 21.4 por ciento por otras personas o incluso por nadie (se quedaron solos).

Se estima que la demanda potencial real por servicios de guarderías en 2005 es de alrededor de 635 mil niños, la capacidad instalada del Instituto cubre tan sólo la tercera parte de esa cifra; lo cual quiere decir que existen únicamente alrededor de 200 mil sitios disponibles. Si en lugar de tomar como referencia la demanda real se tomara como referencia la demanda potencial máxima, la cobertura por parte del IMSS sería ligeramente superior a la quinta parte de esa demanda.

La tesis se presenta a en tres capítulos:

En el capítulo I se presenta el panorama del muestreo utilizado para el presente estudio, mostrando los distintos conceptos del muestreo probabilístico; explicando las ventajas que ofrece cada método elegido, la forma de generar los estimadores de los parámetros y las propiedades distribucionales del estimador. Cada estrategia se evalúa considerando los posibles valores del estimador.

En el capítulo II se analiza la situación de la oferta y la demanda del seguro de guarderías, haciendo un estudio para elegir el diseño muestral más adecuado utilizando las herramientas indispensables para el cálculo del tamaño de muestra, dependiendo de las estimaciones dadas por cada modelo, adaptándola a las necesidades primordiales de la cobertura de seguros de guarderías, empleando las bases estadísticas de la Encuesta Nacional de Empleo y Seguridad Social, Encuesta Nacional de Salud e INEGI.

El capítulo III toma en consideración los resultados del capítulo II y en él se efectúa el análisis estadístico con la información obtenida, tomando en consideración las tasas de participación por ramo de aseguramiento en el uso del servicio de guarderías. Previo al análisis se proporciona una descripción y análisis de los datos, dando un informe detallado de los resultados.

CAPÍTULO I. ELEMENTOS TEORICOS Y TÉCNICOS DEL MUESTREO ESTADÍSTICO

1.1 Introducción

Una condición fundamental para realizar un estudio estadístico de cualquier tipo es considerar que los resultados que se obtengan sean confiables. Las técnicas del muestro se utilizan frecuentemente cuando no es posible contar o medir todos los elementos de la población debido al alto costo que esto implica y a la imposibilidad de acceder a la misma.

El uso del muestreo se extiende al estudio de diversos fenómenos sociales, económicos y políticos. Los planificadores y evaluadores de programas educativos, económicos y de salud pública recurren hoy en día al muestreo como instrumento fundamental.

1.2 Conceptos Generales y Notación

1.2.1 Población.

Una población está determinada por las características que la definen y que son del interés de la investigación; por lo tanto, el conjunto de elementos que poseen esta característica se denomina población o universo. Población es la totalidad de elementos donde se observa el fenómeno a estudiar y donde las unidades de población poseen una característica común, la que se estudia y da origen a los datos de la investigación. Para obtenerla, es necesario medir o contar en cada unidad una o varias características, o clasificar sus unidades de acuerdo con ellas. A partir de los resultados podríamos llegar al conocimiento de unos valores como la media, el total, varianza, etcétera. A los que denominaremos parámetros o características poblacionales.

En este estudió la población de investigación son los niños que tienen derecho a la prestación de guarderías por parte de sus padres que cotizan al Instituto Mexicano del Seguro Social.

1.2.2 Muestra

Es un conjunto de unidades, una porción del total, que representa la conducta similar al universo en su conjunto; o bien, es el conjunto de elementos de la población seleccionados a partir del marco de muestreo. Cuando seleccionamos algunos elementos con la intención de averiguar algo sobre una población determinada, nos referimos a este grupo de elementos como *muestra.*

Las condiciones que debe reunir una muestra son: i) Homogeneidad, que debe ser extraída de la misma población; ii) Independencia, las observaciones no deben estar mutuamente condicionadas entre sí; y iii) Representatividad, la muestra debe ser el mejor reflejo posible del conjunto del cual proviene.

Existen diversas formas de tomar una muestra: i) No probabilística[1], y ii) Probabilística[2] donde todos los elementos de la población tienen una probabilidad conocida y mayor que cero de ser seleccionados.

Entonces, cuando no es posible medir cada uno de los individuos de una población, se toma una muestra representativa de la misma. Para elegir las unidades que constituyen la muestra, es necesario disponer de un conjunto real de unidades, que de represente lo mejor posible al conjunto ideal que constituye la población.

A cada unidad de la población $\{U_i; i = 1,\ldots,N\}$, le haremos corresponder una variable cuantitativa resultado de medir una característica de la misma; o bien una variable cualitativa que tome valores 1 y 0 de acuerdo con la pertenencia, o no, de una característica determinada.

El método de muestreo se basa en ciertas leyes que le otorgan su fundamento científico, las cuales son:

1. Ley de los grandes números: si en una prueba, la probabilidad de un acontecimiento o suceso es P, y si éste se repite una gran cantidad de veces, la relación entre las veces que se produce el suceso y la cantidad total de pruebas (es decir, la frecuencia F del suceso) tiende a acercarse cada vez más a la probabilidad P.
2. Cálculo de probabilidades: La probabilidad de un hecho o suceso es la relación entre el número de casos favorables (p) a este hecho con la cantidad de casos posibles; suponiendo que todos los casos son igualmente posibles. El método de establecer la probabilidad es lo que se denomina cálculo de probabilidad.

De estas dos leyes fundamentales de la estadística, se infieren aquellas que sirven de base al método de muestreo:

---

[1] Estas formas pueden sé *A juicio,* donde se usa la experiencia del investigador; *Cuotas* que consiste en predeterminar la cantidad de elementos de cada categoría que habrá de integrar la muestra; *Muestra accidental*, que se obtiene sin ningún plan preconcebido; las unidades elegidas resultan producto de circunstancias fortuitas; *Muestra intencional* donde las unidades se eligen en forma arbitraria, designando a cada unidad según características que para el investigador resulten de relevancia. En estos casos *puede resultar una muestra sesgada* y *No hay forma de estimar el error*

[2] Hay forma de estimar el error y se tiene apoyo de herramientas de probabilidad.

a. Ley de la regularidad estadística: un conjunto de n unidades tomadas al azar de un conjunto N, puede que tenga las características del grupo más grande.
b. Ley de la inercia de los grandes números: Se refiere al hecho de que en la mayoría de los fenómenos, cuando una parte varía en una dirección, es probable que una parte igual del mismo grupo, varíe en dirección opuesta.
c. Ley de la permanencia de los números pequeños: si una muestra suficientemente grande de la población, una segunda muestra de igual magnitud deberá ser semejante a la primera; y si en la primera muestra se encuentran pocos individuos con características raras, es de esperar encontrar igual proporción en la segunda muestra.

Las situaciones típicas en el empleo de muestras, son las siguientes:

- Dado que la población de los niños de 0-4 años que tienen derecho a disfrutar de los servicios de guardería en el IMSS en el Distrito Federal es de difícil captación, se ha optado por observar sólo una proporción y realizar un muestreo de dicha población; máxime cuando el objeto de este estudio es el análisis de las variables que son determinantes en la cobertura de las guarderías a nivel nacional.
- Las familias que requieren el servicio de guarderías en el Distrito Federal mantienen características comunes, aún cuando difieran en aspectos como el nivel socioeconómico o la región geográfica en la que se encuentran, por lo que tomar una muestra de dicha población nos proporciona una buena representación de la misma.

Por otra parte se pueden mencionar entre las principales ventajas del muestreo, las siguientes:

- Costos: Tomar solo una parte de las familias que requieren el servicio de guarderías reduce en gran parte los costos de medición de las características que proporcionan un perfil de ellas, en comparación de un estudio sobre toda la población en cuestión.
- Precisión: El muestreo proporciona un mejor control de los errores en la recolección de datos que el que proporcionaría un censo, porque la muestra es una representación a menor escala del objetivo.
- Oportunidad: Otra ventaja de una muestra sobre un censo es que la muestra proporciona la información mucho más rápido. Este factor es importante para la toma de decisiones oportunas.
- Calidad de Información: Una muestra puede proporcionar información más detallada que un censo porque recolectar esta toma menos tiempo, es menos costosa y nos permite ser más cuidadosos en cada etapa del procesamiento de la información.

Así mismo es conveniente señalar las principales limitaciones:

- Incertidumbre: riesgo que tienen las muestras de sufrir desvíos o sesgos debido a metodologías o prácticas de campo inadecuadas.

1.2.3 Variables

Para realizar un diseño muestral es necesario considerar el comportamiento de las variables en la población, es por ello que se dan definiciones escenciales de algunos tipos de fenómemos que deseamos medir.

Las variables se pueden definir como todo aquello que vamos a medir, controlar y estudiar en una investigación o estudio. Por lo tanto, es importante, antes de iniciar una investigación, que sepamos cuáles son las variables que vamos a medir y la manera en que lo haremos. Es decir, las variables deben ser susceptibles de medición, las cuales pueden ser definidas conceptual y operacionalmente. [3]

Variable es todo aquello que puede asumir diferentes valores, desde el punto de vista cuantitativo o cualitativo.

Para definir las variables, podemos utilizar los indicadores[4], que constituyen el conjunto de actividades o características propias de un concepto. Por ejemplo, si analizamos la estancia o no de un niño en una guardería, podemos decir que está compuesta por una serie de factores como el nivel socioeconómico de la familia, la situación laboral de ambos padres, etc. Cada factor puede ser medido a través de indicadores. En otras palabras, los indicadores son algo específico y concreto que representan algo más abstracto o difícil de precisar.

No todos los indicadores tienen el mismo valor. Es decir, aunque haya varios indicadores para un mismo fenómeno, habrá algunos más importantes que otros y, por lo general, cualquier indicador que se tenga está basado en una probabilidad de que realmente represente al fenómeno.

[3]. La definición conceptual es de índole teórica, mientras que la operacional nos da las bases de medición y la definición de los indicadores.

[4] Algunos criterios para escoger los indicadores:
- Se debe tener el menor número de indicadores de una variable, siempre y cuando éstos sean realmente representativos de la misma.
- Se deben poseer formas de medición específicas para cada indicador.
- Hay que tener en cuenta que los indicadores sólo poseen una relación de probabilidad con respecto a la variable.

1.2.4 Parámetros y Estimadores

Es necesario explicar algunas definiciones para partir de un lenguaje común y tener claro qué es un parámetro y un estimador, que es base esencial para realizar la estrategia de muestreo que se muestra a continuación.

Una población queda caracterizada a través de ciertos valores denominados **parámetros**, que describen las principales propiedades del conjunto.

Un parámetro es un valor fijo (no aleatorio) que caracteriza a una población en particular. En general, un parámetro es una cantidad desconocida y rara vez se puede determinar exactamente su valor, por la dificultad práctica de observar todas las unidades de una población. Por este motivo, tratamos de estimar el valor de los parámetros desconocidos a través del empleo de muestras. Las cantidades usadas para describir una muestra se denominan **estimadores o estadísticos muéstrales.**

Ahora bien, es razonable pensar que si tomamos diferentes muestras de la misma población y calculamos los diferentes estadísticos de cada una, esos valores van a diferir de muestra a muestra. Por lo tanto, un estadístico no es un valor fijo, sino que presenta las siguientes características:

- Puede tener varios resultados posibles.
- No se puede predecir de antemano su valor.

Estas son las condiciones que definen a una variable aleatoria. Un estadístico, entonces, es una variable aleatoria, función de las observaciones muéstrales.

Si un estadístico es una variable aleatoria, entonces es posible determinar su distribución de probabilidades y calcular sus principales propiedades.

Los pará

metros o características poblaciones pueden expresarse con:

N El total de la población

$\theta = \sum_{i=1}^{N} Y_i$ que para valores particulares nos da:

$Y_i = X_i$ $\qquad \theta = \sum_{i=1}^{N} X_i$ (total)

$Y_i = \frac{X_i}{N}$ $\qquad \theta = \sum_{i=1}^{N} \frac{X_i}{N} = \frac{X}{N}$ (media)

$$Y_i = A_i \qquad \theta = \sum_{i=1}^{N} A_i \qquad \text{(total de clase)}$$

$$Y_i = \frac{A_i}{N} \qquad \theta = \sum_{i=1}^{N} \frac{A_i}{N} = P \qquad \text{(Proporción)}$$

$$Y_i = \frac{(X_i - \overline{X})^2}{N} \qquad \theta = \sum \frac{(X_i - \overline{X})^2}{N} = \sigma^2 \qquad \text{(Varianza)}$$

$$Y_i = \frac{(X_i - \overline{X})^2}{N-1} \qquad \theta = \sum_{i=1}^{N} \frac{(X_i - \overline{X})^2}{N-1} = s^2 \qquad \text{(Varianza ajustada)}$$

El conjunto de párametros poblacionales constituyen el espacio métrico para Ω.

1.3 Panorama del Muestreo

Las encuestas por muestreo son un tipo de investigaciones que tienen como propósito conocer algo respecto a una determinada población humana, y estudian solo una parte de ésta,[5] también denominada Demoscopía (la investigación de la población), es decir, la disciplina o grupo de ellas que pretende conocer algún aspecto de una población o conjunto de seres humanos.[6]

El enfoque científico consiste básicamente en usar los conocimientos previos que se tienen sobre el problema y diseñar una metodología de investigación que minimice la ocurrencia y magnitud de los errores.

1.3.1 Marco de Muestreo

A la definición conceptual de población objetivo debe corresponder una forma práctica de tener acceso a todos los elementos que la constituyen. En forma simple, el marco de muestreo es el medio físico en el que se identifica a todos los elementos de una población; en nuestro caso el marco de muestreo son los directorios proporcionados por IMSS, donde se encuentra ubicadas las unidades informantes (guarderías). Es importante señalar que el marco de muestreo es la base "necesaria" en el proceso de selección de la muestra y si no se cuenta con éste el proceso simplemente no se puede llevar a cabo.

La disposición de marcos adecuados puede ser una problemática frecuente, muchas veces los archivos documentales disponibles que usualmente son

[5] Para efectos de este estudio nos referimos a niños de 0-4 años.
[6] Las formas de obtener información en la Demoscopía son a través de censos, registros administrativos y encuestas por muestreo.

utilizados como marco de muestreo suelen presentar omisiones de elementos de la "población objetivo", repeticiones o inclusiones de elementos extraños, dando como resultado diversos escenarios, por ejemplo el escenario "ideal", el marco que incluye otros elementos adicionales[7], el marco que no cubre todos los elementos[8], marcos complementarios[9] y marcos traslapados[10].

En nuestro caso contamos con el marco ideal, esto porque el marco y la población coinciden, es decir, no existe ningún problema para su utilización en el proceso de selección de la muestra.

1.3.2 Representatividad de la Muestra

Cuando seleccionamos algunos elementos con la intención de averiguar algo sobre una población determinada, por supuesto, esperamos que lo que averiguamos en la muestra sea cierto para la población en su conjunto. Cuando una muestra cumple con esta condición, es decir, cuando nos refleja en sus unidades lo que ocurre en el universo, la llamamos muestra **representativa**. Por lo tanto, una *muestra representativa* contiene las características relevantes de la población en las mismas proporciones en que están incluidas en tal población. Sus conclusiones son susceptibles de ser generalizadas al conjunto del universo, aunque para ello debamos añadir un cierto margen de error en nuestras estimaciones.

La muestra descansa en el principio de que las partes representan al todo y, por tal, refleja las características que definen la población de la que fue extraída, lo cual nos indica que es representativa. Por lo tanto, la validez de la generalización depende de la validez y tamaño de la muestra.

1.3.3 Teorema Central del Límite

Un teorema fundamental en estadística, que es de suma importancia, es el Teorema Central del Límite, y que dice:

- La media de la distribución de muestreo será igual a la media de la población.
- Al incrementarse el tamaño de la muestra, la distribución de muestreo de la media se acercará a la normalidad, sin importar la forma de la distribución de la población.

---

[7]Donde es necesario desechar aquellos elementos que no se consideran miembros.

[8] Se puede optar por redefinir la población de manera que coincida con el marco ó complementar el marco con otros marcos.

[9] Cuando la unión de dos o mas marcos cubre a toda la población, cuando esto sucede se deberá usar muestreo con estratos, donde el número mínimo de estratos será igual al número de marcos que forman la unión.

[10] Cuando algunas de las unidades de medición se encuentran en varios marcos, en este caso deben identificarse estas unidades y ser asignadas a solo uno de ellos.

Esta relación entre la forma de la distribución de la población y la forma de la distribución de muestreo se denomina *teorema central del límite*, que es tal vez el más importante de toda la inferencia estadística. Nos asegura que la distribución de muestreo de la media se aproxima a la normal al incrementarse el tamaño de la muestra. De manera laxa, dice que los promedios de *muchas muestras* probabilísticas de una población tienden, al aumentar de tamaño de muestra *n,* a tener una distribución normal, a pesar de que la variable que se mida no tenga una distribución normal en la población. Una muestra no tiene que ser muy grande para que la distribución de muestreo de la media se acerque a la normal. Los estadísticos utilizan la distribución normal como una aproximación a la distribución de muestreo siempre que el tamaño de la muestra sea al menos de 30, pero la distribución de muestreo de la media puede ser casi normal con muestras incluso de la mitad de ese tamaño. La importancia del teorema central del límite es que nos permite usar estadísticas de muestra para hacer inferencias con respecto a los parámetros de población sin saber nada sobre la forma de la distribución de frecuencias de esa población más que lo que podamos obtener de la muestra.

Sea $y_1, y_2, \ldots, y_n$ una muestra aleatoria de una densidad de probabilidad $f_Y(y)$ con media μ y varianza σ^2. Sea $\overline{y} = \dfrac{y_1 + y_2 + \ldots + y_n}{n}$ la media muestral para un tamaño de muestra *n* ***grande*** la distribución de y converge a la distribución normal se tiene que:

$$\overline{y} \approx N\left(\mu, \frac{\sigma^2}{n}\right)$$

ó bien $\lim_{n\to\infty} P\left[a < \dfrac{\overline{y}-\mu}{\frac{\sigma}{\sqrt{n}}} \le b\right] = \dfrac{1}{\sqrt{2\pi}} \int_a^b e^{\frac{-x^2}{2}}\, dx \qquad \forall a,b \qquad a < b$

A $\sqrt{\dfrac{\sigma^2}{n}} = \dfrac{\sigma}{\sqrt{n}}$ se llama "error estándar" de $\overline{Y}$, y es la distribución estándar de la distribución de los posibles valores de $\overline{Y}$.

En el teorema Central del Límite es importante señalar que si la muestra se toma de una población infinita (o finita con reemplazo), los valores de Y_i son variables aleatorias independientes e idénticamente distribuidas.

1.3.4 Estrategias del Muestreo y propiedades Distribucionales

Para poder realizar un muestreo, es necesario especificar, la forma de tomar el tamaño de la muestra y calcular las estimaciones del parámetro $\hat{\theta}$ de interés, a esto se le llama estrategia de muestreo.

Para determinar el tamaño de muestra es necesario tomar en cuenta la variabilidad de lo que se quiere estudiar, la precisión con que se quiere hacer la inferencia, el presupuesto disponible y el tamaño de la población.

La distribución de probabilidad del estimador $\hat{\theta}$ sobre el espacio muestral se denomina "distribución del estimador en el muestreo". El estimador es, por consiguiente, una variable aleatoria cuyos valores particulares son las estimaciones. Para cada estrategia de muestreo, que comprende el diseño y el estimador, se determinan las propiedades de distribución de las posibles muestras.

Para llevar acabo las estimaciones del parámetro es importante que sea insesgado, es decir, $E(\hat{\theta}) = \theta$, para determinar cuál estimador es *mejor* entre varios posibles. Para un tamaño de muestra fijo, el mejor estimador es el que tenga *menor variabilidad alrededor de su media* (o sea el parámetro, si el estimador es insesgado). Si por el contrario el estimador es sesgado, es necesario que este sea consistente, es decir, que el límite de $\hat{\theta}$ cuando $n \to \infty$ sea igual al parámetro θ de interés.

El espacio muestral S(x), la función de probabilidad P(x) y el estimador $\hat{\theta}(x)$, constituyen el "diseño muestral":

$$D = \{ S(x) \ ; \ P(x) \ ; \ \hat{\theta}(x) \}$$

La variabilidad en las estadísticas de muestras proviene del *error de muestreo* debido al azar; es decir, hay diferencias entre cada muestra y la población, y entre las diversas muestras, debido únicamente a los elementos que decidimos escoger para las muestras.

La medida de variabilidad es la varianza del estimador:

Si el estimador no es insesgado es conveniente tomar como medida de variabilidad el error cuadrático medio (E.C.M), de esto, se deduce que en el error debido al muestreo existe una componente aleatoria o "error de muestreo" que se mide por la desviación típica del estimador, y puede existir una componente sistemática o sesgo ($B = E(\hat{\theta}) - \hat{\theta}$).

Al tener muchos valores de un estimador se puede estudiar su regularidad estadística con su distribución de frecuencias. Estamos hablando de una nueva población, que son las diferentes estimaciones del parámetro de interés.

Se pueden manejar varias estrategias de muestreo cuya selección depende de los objetivos del estudio, del tiempo límite y de los recursos disponibles.

El muestreo debe realizarse sobre una base estadística. La estrategia de muestreo debe garantizar la representatividad de la muestra de la población analizada y la eficacia del método de muestreo.
Se deben tener en cuenta los siguientes criterios:

- tamaño de la muestra;
- origen de la muestra;
- especie;
- categoría dentro de la especie;
- estratificación dentro de la categoría;
- muestra aleatoria (específica, sistemática);
- tipo de muestra;

1.4 El Muestreo Probabilístico

Los muestreos probabilísticos son métodos que utilizan un proceso que garantiza que puede calcularse la probabilidad de extracción de cualquiera de las muestras posibles. Este método permite calcular la representatividad o no de una muestra.

Las técnicas de *muestreo probabilístico* son aquellas en las que se determina aleatoriamente (con el fin de evitar sesgos) a los individuos que constituirán la muestra. Estas técnicas nos sirven cuando se desean generalizar los resultados que se obtienen a partir de la muestra hacia toda la población. Lo anterior se dice dado que supone que el proceso aleatorio permitirá la obtención de una muestra *representativa* de la población.

Cuando el método de muestreo define, para un conjunto de especificaciones, una función de probabilidad P(x) tal que: $\sum_{x \in S(x)} P(x) = 1$, y un método de estimación que proporcione para cada muestra una estimación, el muestreo se denomina "muestreo probabilístico". Si definimos el conjunto de muestras distintas S_1, S_2, ... , S_v, donde el procedimiento es capaz de elegir, si se aplica a una población específica. Esto significa que podemos decir con precisión cuáles son las unidades de muestreo que pertenecen a S_1, S_2, etc., entonces cada muestra posible S_i que se selecciona por un proceso aleatorio, que es una aplicación del muestreo de probabilidad según el cual cada unidad de la población tiene la misma oportunidad de ser incluido en la muestra, tiene una probabilidad de selección π_i de ser elegida. El método de muestreo para calcular la estimación a partir de la muestra debe ser definido y conducir a una estimación única para cualquier muestra específica. Podemos decir, por ejemplo, que las estimaciones es el promedio de las mediciones correspondientes a las unidades individuales de la muestra.

Para todo procedimiento de muestreo que satisfaga lo anterior, podemos calcular la distribución de frecuencia de las estimaciones que genera el proceso, si se aplica repetidamente a la misma población, donde la frecuencia con que se elige cualquier muestra S_i, y se sabe como calcular la estimación a partir de los datos de S_i, entonces se puede desarrollar una teoría de muestreo para cada procedimiento de este tipo, en la práctica rara vez se extrae una muestra de probabilidad dando las S_i y los números π_i, por lo general, la extracción se hace al especificar la probabilidades de inclusión en la muestra para las unidades individuales y extraer unidades, una a la vez, o en grupos, hasta constituir muestra del tamaño y tipo deseado.

Con este método se confirma, como se menciono anteriormente, con la ayuda de la teoría de muestreo y la distribución normal es posible predecir, en forma aproximada y con los datos de la muestra, la magnitud del error esperado en las estimaciones hechas a partir de la muestra.

1.5 Muestreo Aleatorio Simple (m.a.s)

Es un método de selección de n unidades en un conjunto de N de tal modo que cada de las una de la ${}_N C_n$ muestras distintas, tengan la misma oportunidad de ser elegidas. Se enumeran las unidades de 1 a N. posteriormente se extrae una serie de n números aleatorios entre 1 y N, ya sea utilizando una tabla de números aleatorios o mediante un programa de computación que produce una tabla semejante. Las unidades que llevan estos números constituyen la muestra.

Este esquema de muestreo es el más usado cuando se tiene un marco de muestreo que especifique la manera de *identificar* cada unidad en la población.

Se asume que no se tiene conocimiento *a priori* sobre los posibles valores.

1.5.1 Muestreo aleatorio simple con reemplazo

Los *muestreos con reemplazo* son aquellos en los que una vez que ha sido seleccionado un individuo (y estudiado) se le toma en cuenta nuevamente al elegir el siguiente individuo a ser estudiado. En este caso cada una de las observaciones permanece independiente de las demás, pero con poblaciones pequeñas (un grupo de escuela de 30 alumnos, por ejemplo) tal procedimiento debe ser considerado ante la posibilidad de repetir observaciones. En el caso de poblaciones grandes no importa tal proceder, pues una repetición no afecta sustancialmente a las frecuencias relativas.

Si cada unidad se toma con reemplazo, la probabilidad de una muestra particular es:

$$\frac{1}{N}\cdot\frac{1}{N}\cdots\frac{1}{N}=\frac{1}{N^n}$$

1.5.2 Muestreo aleatorio simple sin reemplazo

Los *muestreos sin reemplazo* son los que una vez que se ha tomado en cuenta un individuo para formar parte de la muestra, no se le vuelve a tomar en cuenta nuevamente. En este caso, y hablando específicamente para el caso de poblaciones pequeñas, las observaciones son dependientes entre sí, pues al no tomar en cuenta nuevamente el individuo se altera la probabilidad para la selección de otro individuo de la población. Para el caso de las poblaciones grandes (por ejemplo la población de un país) dicha probabilidad para la selección de un individuo se mantiene prácticamente igual, por lo que se puede decir que existe independencia en las observaciones.

1.5.2.1 Definición y notación

Se tiene una población de N unidades y se toma una muestra con las siguientes características:

- El tamaño de la muestra es n.
- Las unidades se seleccionan sin reemplazo, lo que equivale a selecciones sucesivas e independientes con probabilidades para cada extracción iguales a: $\frac{1}{N-i}$ $\quad \forall\, i = 0, 1, 2, \ldots, n\text{-}1$

Las muestras que tengan las mismas unidades aunque el orden de extracción sean distintos se consideran iguales.

Puesto que se seleccionan sin reemplazo y el orden no importa, el número total de muestra está dado por todas las formas posibles de seleccionar n unidades de N en total. Este número de formas corresponden a las combinaciones:

$$\binom{N}{n}=\frac{N!}{(N-n)!\,n!}$$

Notación:
N Tamaño de la población
n Tamaño de la muestra

y_i El valor de la variable estudiada en la i-ésima unidad de la muestra o de la población

Y Total de la población

$$Y = \sum_{i=1}^{N} y_i$$

$\overline{Y}$ Media de la población

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

$\frac{n}{N}$ Fracción del muestreo

$\hat{Y}$ Estimador del total

$\hat{\overline{Y}}$ Estimador de la media

$\overline{y}$ Media de la muestra

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

Una muestra aleatoria donde cualquier elemento $U_{i,j} = 1,\ldots,N$ tiene probabilidad $\frac{1}{\binom{N}{n}}$ se verifica de la siguiente manera: si las unidades de una muestra particular toman valores $y_1, y_2, \ldots, y_n$ la probabilidad de obtenerlas es:

$$\frac{1}{N} \cdot \frac{1}{N-1} \cdot \frac{1}{N-2} \cdots \frac{1}{N-n+1} = \frac{(N-n)!}{N!}$$

Como no importa el orden, se multiplica por todas las posibles formas de ordenar n elementos tomados a la vez, es decir $n!$

$$\frac{(N-n)!\,n!}{N!} = \frac{1}{\binom{N}{n}}$$

En un muestreo aleatorio simple sin reemplazo, la probabilidad de seleccionar el primer elemento de la muestra, esta dada por: $\frac{1}{N}$, es decir, hacer la primer extracción.

En la segunda extracción, su probabilidad de ser seleccionada es $\frac{1}{N-1}$, condicionada a la probabilidad de no haber sido seleccionada en la primera extracción $\frac{N-1}{N}$, entonces tenemos que: $\frac{N-1}{N} \cdot \frac{1}{N-1} = \frac{1}{N}$

En la tercera extracción, la probabilidad de selección es $\frac{1}{N-2}$, condicionada a la probabilidad de no haber sido seleccionada ni en la primera como en la segunda extracción $\left(\frac{N-1}{N}\right)\left(1-\frac{N-1}{N}\right)$, entonces la probabilidad de la tercera elección es:

$$\frac{N-1}{N}\cdot\frac{N-2}{N-1}\cdot\frac{1}{N-2}=\frac{1}{N}$$

La probabilidad de selección de n-ésima de la unidad, tenemos que:

$$\frac{N-1}{N}\cdot\frac{N-2}{N-1}\cdots\frac{1}{N-(n-1)}=\frac{1}{N}$$

Como los eventos son excluyentes, la probabilidad esta dada por la suma de probabilidades, entonces la probabilidad de observar la unidad en la $1^a, 2^a$, o n-ésima extracción es:

$$\frac{1}{N}+\frac{1}{N}+\ldots+\frac{1}{N}=\frac{n}{N}$$

Donde llamaremos la probabilidad de inclusión de primer orden a:

$\pi_j=\frac{n}{N}$, y al inverso factor de expansión: $\frac{1}{\pi_j}=\frac{N}{n}$

1.5.2.2 Estimadores y varianza

Mediante el proceso de muestreo lo que se desea es hacer una inferencia sobre una población; específicamente se desea calcular una estimación de un parámetro de la población.

El estimador más conveniente de la media poblacional $\bar{Y}$ es la media muestral.

$\hat{\bar{Y}}=y$, se probará que $\bar{y}$ es un estimador insesgado $\bar{Y}$.

Por demostrar que $E(\bar{y})=\bar{Y}$

Prueba: Por definición de esperanza e hipótesis de aplicación del muestreo aleatorio simple

$$E(\bar{y})=\frac{\sum_{I=1}^{\binom{N}{n}}\bar{y}_I}{\binom{N}{n}}=\frac{\sum_{i=1}^{\binom{N}{n}}(y_1+y_2+\ldots+y_n)}{nN!\big/n!(N-n)!}$$

Donde la suma se puede calcular siempre y cuando se conozca el número de muestras en las que aparece una particular y_i. Teniendo fija la y_i, *N-1* unidades posibles para asignar en *n-1* lugares, por lo que el número de las muestras en las que aparece una unidad en particular es:

$$\binom{N-1}{n-1} = \frac{(N-1)!}{(n-1)!\,(N-n)!}$$

Esto implica que la suma equivale a:

$$\sum_{i=1}^{\binom{N}{n}} (y_1 + y_2 + \ldots + y_n) = \frac{(N-1)!}{(n-1)!\,(N-n)!}(y_1 + y_2 + \ldots + y_N),$$

Desarrollando obtenemos que:

$$E(\bar{y}) = \frac{\sum_{i=1}^{\binom{N}{n}} (y_1 + y_2 + \ldots + y_n)}{\frac{nN!}{n!\,(N-n)!}} = \frac{n!\,(N-1)(N-n)!}{(n-1)!\,(N-n)!}(y_1 + y_2 + \ldots + y_N) = \frac{y_1 + y_2 + \ldots + y_n}{N} = \bar{Y}$$

Por lo tanto $E(\bar{y}) = \bar{Y}$ es insesgado

Para ver el estimador insesgado de la varianza, se analizan los siguientes resultados.

Se define como:

$$Z_i \begin{cases} 1 & U_i \quad \textit{está en la muestra} \\ 0 & U_i \quad \textit{no está en la muestra} \end{cases} \qquad i = 1, \ldots, N$$

$\{Z_1, Z_2, \ldots, Z_n\}$ variables aleatorias Bernoulli $\left(\frac{n}{N}\right)$, es decir,

$$\left\{ Bernoulli \quad f(x/\pi) = \pi^x (1-\pi)^{1-x} \qquad \begin{matrix} 0 \le \pi \le 1 \\ x = 0,1 \end{matrix} \right\}$$

Calculando su esperanza y varianza a través de su función de densidad tenemos que:

$$E(Z_i) = 0(1-\pi) + 1(\pi) = \pi = \frac{n}{N}$$

esperanza

$$E(Z_i^2) = 0^2(1-\pi) + 1^2(\pi) = \pi = \frac{n}{N}$$

$$V(Z_i) = E(Z_i^2) - E(Z_i)^2 = \pi - \pi^2 = \pi(1-\pi) = \frac{n}{N}\left(1 - \frac{n}{N}\right)$$

Para calcular la Covarianza primero veamos que:
Para $i \neq j$

$$E(Z_i Z_j) = P\lfloor Z_i = 1 \quad y \quad Z_j = 1 \rfloor = P\lfloor Z_j = 1 / Z_i = 1 \rfloor P[Z_i = 1]$$

$$= \left(\frac{n-1}{N-1}\right)\left(\frac{n}{n}\right)$$

Entonces:
Para $i \neq j$

$$Cov\ (Z_i Z_j) = E(Z_i Z_j) - E(Z_i)E(Z_j) = \left(\frac{n-1}{N-1}\right)\left(\frac{n}{N}\right) - \left(\frac{n}{N}\right)^2$$

$$= \left(\frac{n^2 - n}{N^2 - N}\right) - \left(\frac{n^2}{N^2}\right) = \frac{n}{N}\left(\frac{n-1}{N-1} - \frac{n}{N}\right) = \frac{n}{N}\left(\frac{N(n-1) - n(N-1)}{N(N-1)}\right)$$

$$= \frac{n}{N}\left(\frac{Nn - N - nN + n}{N(N-1)}\right) = \frac{n}{N}\left(\frac{n-N}{N(N-1)}\right) = -\frac{1}{N-1}\left(1 - \frac{n}{N}\right)\frac{n}{N}$$

Por lo tanto No son independientes

Estimemos a la varianza:

$$V(\bar{y}) = V\left(\frac{1}{n}\sum_{i=1}^{N} Z_i y_i\right)$$

$$= \frac{1}{n^2}\left[\sum_{i=1}^{N} y_i^2\ V(Z_i) + \sum_{i=1}^{N}\sum_{j \neq i}^{N} y_i y_j\ Cov\ (Z_i, Z_j)\right]$$

$$= \frac{1}{n^2}\left[\frac{n}{N}\left(1 - \frac{n}{N}\right)\sum_{i=1}^{N} y_i^2 - \frac{1}{N-1}\left(1 - \frac{n}{N}\right)\left(\frac{n}{N}\right)\sum_{i=1}^{N}\sum_{i \neq j}^{N} y_i y_j\right]$$

$$= \frac{1}{n^2}\ \frac{n}{N}\left(1 - \frac{n}{N}\right)\left[\sum_{i=1}^{N} y_i^2 - \frac{1}{N-1}\sum_{i=1}^{N}\sum_{i \neq j}^{N} y_i y_j\right]$$

$$= \frac{1}{n}\left(1 - \frac{n}{N}\right)\frac{1}{N(N-1)}\left[(N-1)\sum_{i=1}^{N} y_i^2 - \left(\sum_{i=1}^{N} y_i\right)^2 + \sum_{i=1}^{N} y_i^2\right]$$

Sabemos que:

$$\left(\sum_{i=1}^{N} y_i\right)^2 = \left[\sum_{i=1}^{N} y_i^2 + \sum_{i=1}^{N}\sum_{i \neq j}^{N} y_i y_j\right]$$

$$S^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(y_i - \overline{Y}\right)^2 = \frac{1}{N-1}\left[\sum_{i=1}^{N} y_i^{\,2} - N\overline{Y}^{\,2}\right]$$

Entonces:

$$V(\overline{y}) = \frac{1}{n}\left(1-\frac{n}{N}\right)\frac{1}{N(N-1)}\left[N\sum_{i=1}^{N} y_i^{\,2} - \left(\sum_{i=1}^{N} y_i\right)^2\right]$$

$$= \frac{1}{n}\left(1-\frac{n}{N}\right)\frac{1}{N(N-1)}\,N\left[\sum_{i=1}^{N} y_i^{\,2} - N\overline{Y}^{\,2}\right]$$

$$= \left(1-\frac{n}{N}\right)\frac{S^2}{n} = V(\overline{y})$$

$V(y)$ se estima insesgadamente con: $\hat{V}(\overline{y}) = \left(1-\frac{n}{N}\right)\frac{\hat{S}^2}{n}$

Para mostrar $\hat{V}(\overline{y}) = \left(1-\frac{n}{N}\right)\frac{\hat{S}^2}{n}$ es un estimador insesgado de $V(\overline{y}) = \left(1-\frac{n}{N}\right)\frac{S^2}{n}$, solo basta demostrar que: $E\left(\hat{S}^2\right) = S^2$

Sabemos que: $S^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(Y_i - \overline{Y}\right)^2$; $\hat{S}^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(y_i - \overline{y}\right)^2$

Entonces:

$$E\left[\frac{1}{n-1}\sum_{i=1}^{n}\left(y_i - \overline{y}\right)^2\right] = \frac{1}{n-1}E\left[\sum_{i=1}^{n}\left\{(y_1 - \overline{Y}) - (\overline{y} - \overline{Y})\right\}^2\right]$$

$$= \frac{1}{n-1}E\left[\sum_{i=1}^{n}(y_i - \overline{Y})^2 - 2\sum_{i=1}^{n}(y_i - \overline{Y})(\overline{y} - \overline{Y}) + \sum_{i=1}^{n}(\overline{y} - \overline{Y})^2\right]$$

$$= \frac{1}{n-1}E\left[\sum_{i=1}^{n} Z_i\,(y_i - \overline{Y})^2 - 2(\overline{y} - \overline{Y})\sum_{i=1}^{n}(y_i - \overline{Y}) + n(\overline{y} - \overline{Y})^2\right]$$

$$= \frac{1}{n-1}E\left[\sum_{i=1}^{n} Z_i\,(y_i - \overline{Y})^2 - 2n(\overline{y} - \overline{Y})^2 + n(\overline{y} - \overline{Y})^2\right]$$

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

$$= \frac{1}{n-1}\left[\frac{n}{N}\sum_{i=1}^{n}(y_i - \overline{Y})^2 - \left(1-\frac{n}{N}\right)S^2\right]$$

$$= \frac{1}{n-1}\left[\frac{n\,(N-1)}{N}S^2 - \left(1-\frac{n}{N}\right)S^2\right]$$

$$= \frac{1}{n-1} S^2 \left[\frac{n\,(N-1)}{N} - \frac{N-n}{N} \right]$$
$$= \frac{1}{n-1} S^2 \, \frac{n(N-1) - N + n}{N}$$
$$= \frac{1}{n-1} S^2 \, \frac{N(n-1)}{N} = S^2$$

Por lo tanto $\hat{V}(\bar{y})$ es un estimador insesgado de $V(\bar{y})$

Es importante recalcar que $V(\bar{y})$ se tiene que estimar, ya que en general no conocemos S^2

Intervalo de Confianza

El intervalo de confianza se refiere a la variación que se puede dar debido a que la muestra elegida es una de las posibles muestras que se pueden producir, Si seleccionamos muchas muestras aleatorias del mismo tamaño es posible calcular un intervalo de confianza para cada una de las muestras, y así tendremos un porcentaje de confianza determinando, donde la media de la población caerá dentro del intervalo. Es decir, el intervalo de confianza es el intervalo de valores que tiene designada una probabilidad de que incluya el valor real del parámetro de la población.

La probabilidad que asociamos con una estimación de intervalo se conoce como nivel de confianza. Esta probabilidad indica qué tanta confianza tenemos de que la estimación de intervalo incluya al parámetro de población. Una probabilidad más alta indica más confianza.

El intervalo de confianza es el alcance de la estimación que estamos haciendo. Expresaremos el intervalo de confianza en términos de desviación estándar, más que con valores numéricos. Los límites de confianza son los límites superior e inferior del intervalo de confianza

Por el Teorema Central de Límite se puede suponer que, con n suficientemente grande, los estimadores de la media $\bar{Y}$ y el total Y se distribuyen en forma normal en torno a los parámetros. Sin embargo, influye de manera definitiva el conocimiento previo que se tenga de la variable, ya que las variables con comportamiento asimétrico requieren de un tamaño mayor para su convergencia a la normalidad que los requeridos para las variables de comportamiento simétrico. Las muestras de poblaciones asimétricas suelen conservar esta asimetría en las distribución de sus correspondientes medias, es por eso que se supone que:

$\bar{y} \sim N\left(\bar{Y}, V(\bar{y})\right)$ ó bien si estandarizamos tenemos que:

$$\frac{\bar{y}-\bar{Y}}{\sqrt{V(\bar{y})}} \sim \ N(0,1)$$

Un intervalo del $100(1-\alpha)\%$ de confianza esta dado por:
Si *Z→(0,1)*
$P\left(-Z_{(1-\alpha/2)} < Z < Z_{(1-\alpha/2)}\right) = 1-\alpha$

Donde $Z_{(1-\alpha/2)}$ es el cuantil de una normal

Tomando la variable *Z*:

$$Z = \frac{\bar{y}-\bar{Y}}{\sqrt{\left(1-\frac{n}{N}\right)\frac{S^2}{n}}}$$

$$\Rightarrow P\left(-Z_{(1-\alpha/2)} < \frac{\bar{y}-\bar{Y}}{\sqrt{\left(1-\frac{n}{N}\right)\frac{S^2}{n}}} < Z_{(1-\alpha/2)}\right) = 1-\alpha$$

$$\Rightarrow P\left(\bar{y} - Z_{(1-\alpha/2)}\sqrt{\left(1-\frac{n}{N}\right)\frac{S^2}{n}} < \bar{Y} < \bar{y} + Z_{(1-\alpha/2)}\sqrt{\left(1-\frac{n}{N}\right)\frac{S^2}{n}}\right) = 1-\alpha$$

De tal manera, los límites del intervalo son:

$$\bar{y} \pm Z_{(1-\alpha/2)}\sqrt{\left(1-\frac{n}{N}\right)\frac{S^2}{n}}$$

Cuando no se conoce $V(\bar{y})$ y se estima con $\hat{V}(\bar{y})$ entonces;

$$\frac{\bar{y}-\bar{Y}}{\sqrt{V(\bar{y})}} \sim t_{n-1}$$

Y el intervalo aproximado de $100(1-\alpha)\%$ de confianza para $\bar{Y}$ es:
$\bar{y} \pm t_{n-1}^{\alpha/2}\sqrt{\hat{V}(\bar{y})}$
En general, como *n* es grande, el valor de la *t* se aproxima a la normal y se usa como intervalo de confianza: $\bar{y} \pm Z_{(1-\alpha/2)}\sqrt{\left(1-\frac{n}{N}\right)\frac{S^2}{n}}$

ESTIMADOR DEL TOTAL

$Y = \sum_{i=1}^{N} Y_i = N\bar{Y}$, ahora bien;

$$\hat{Y} = N\hat{\bar{Y}} = N\bar{y} = \sum_{i=1}^{N}\frac{N}{n}y_i$$

$E(\hat{Y}) = Y$ por lo tanto $\hat{Y}$ es un estimador insesgado de Y

La varianza del estimador total $\hat{Y} = N\overline{y}$ está dada por:

$$V(\hat{Y}) = N^2\left(1 - \frac{n}{N}\right)\frac{S^2}{n}$$

Demostración:

$$\begin{aligned} V(\hat{Y}) &= V(N\overline{y}) \\ &= N^2 V(\overline{y}) \\ &= N^2\left(1 - \frac{n}{N}\right)\frac{S^2}{n} = V(\hat{Y}) \end{aligned}$$

Donde

$$S^2 = s^2 = \frac{\sum_{i=1}^{n}(y_i - \overline{y})^2}{n-1}$$

Debido a que $\hat{Y} = N\overline{y}$, se tiene como corolario que los límites para un intervalo $100(1-\alpha)\%$ para el total de Y es:

$$\hat{Y} \pm NZ_{(1-\alpha/2)}\sqrt{\left(1 - \frac{n}{N}\right)\frac{S^2}{n}}$$

1.5.2.3 Determinación del tamaño de muestra

Siempre que tomamos una muestra perdemos algo de información útil con respecto a la población. El error de muestreo se puede controlar si seleccionamos una muestra cuyo tamaño sea el adecuado. En general, cuanta más precisión se quiera, más grande será el tamaño de la muestra necesaria.

El error de muestreo es una medición de la diferencia entre el parámetro de la población y el valor que toma el estimador para una muestra dada. Al disminuir el error de muestreo, el valor del estimador de una muestra dada probablemente se acercará al valor del parámetro de la población. Los estadísticos describen este fenómeno diciendo: al disminuir el error de muestreo, se incrementa la precisión. Entonces para calcular el tamaño de muestra adecuado es necesario establecer una relación probabilística en torno a un error máximo admisible y así obtener el tamaño de la muestra mínimo.

Para establecer el tamaño de muestra primero fijemos la precisión δ y una confianza $(1-\alpha)$

$$P\left(\left|\overline{y}-\overline{Y}\right|<\delta\right)=1-\alpha$$

$$P\left(\overline{y}-\delta<\overline{Y}<\overline{y}+\delta\right)=1-\alpha$$

Por otro lado sabemos:

$$P\left(-Z_{(1-\alpha/2)}<\frac{\overline{y}-\overline{Y}}{\sqrt{V(\overline{y})}}<Z_{(1-\alpha/2)}\right)=1-\alpha$$

$$P\left(\overline{y}-Z_{(1-\alpha/2)}\sqrt{V(\overline{y})}<\overline{Y}<\overline{y}+Z_{(1-\alpha/2)}\right)=1-\alpha$$

Por lo tanto,

$$\delta=Z_{(1-\alpha/2)}\sqrt{V(\overline{y})}$$

Si n entonces:

$$\delta=Z_{(1-\alpha/2)}\sqrt{\left(1-\frac{n}{N}\right)\frac{S^2}{n}}=Z_{(1-\alpha/2)}\sqrt{\left(\frac{1}{n}-\frac{1}{N}\right)S^2}$$

Despejando n para tener el tamaño de muestra tenemos que:

$$n=\frac{1}{\dfrac{\delta^2}{S^2Z^2_{1-\alpha/2}}+\dfrac{1}{N}}$$

Si N es grande:

$$n_0=\frac{\delta^2}{S^2Z^2_{1-\alpha/2}}$$

Si n no es grande:

$$n=\frac{1}{\dfrac{1}{n_0}+\dfrac{1}{N}}=\frac{n_0}{1+\dfrac{n_0}{N}}$$

δ es el error absoluto

Es importante señalar que se tiene que conocer S^2 para calcular el tamaño de la muestra.

Si se considera el supuesto de normalidad para el estimador,

$$n_0=\frac{S^2N^2Z^2_{1-\alpha/2}}{\delta^2}$$

Donde,

$$n = \begin{cases} n_0 & \text{Si N es grande} \\ \dfrac{n_0}{1+\dfrac{n_0}{N}} & \text{Si N no es grande} \end{cases}$$

1.5.2.4 Estimación de Proporciones

Usualmente se desea estimar a proporción o total de unidades en una población que poseen determinada característica o atributo. $Y(U_i)$ es una medida o indicador de la *presencia o ausencia de una característica* en la unidad U_i con valor 1 si la característica está presente y 0 si no es así. En este caso $\bar{Y}$ =*P*=*proporción* de unidades en la *población* que tienen la característica

Se define Yi_1 como:

$$Yi \begin{cases} 1 \quad U_i \quad \textit{tiene la característica de A} \\ 0 \quad U_i \quad \textit{no tiene la característica de A} \end{cases}$$

$$P = \frac{\text{\# de elementos que tiene la característica de A}}{\text{Total de elementos}} = \frac{\sum_{i=1}^{N} Y_i}{N}$$

Un estimador insesgado de P es:

$$\hat{P} = \frac{\sum_{i=1}^{n} y_i}{n} = \bar{y}$$

Con varianza: $V(\hat{P}) = \left(1 - \dfrac{n}{N}\right)\dfrac{1}{n} \cdot \dfrac{N}{N-1} P(1-P)$

El número de elementos que tiene la característica A es:

$$\sum_{i=1}^{N} Y_i = \sum_{i=1}^{N} Y_i^2 = NP$$

Entonces:

$$V(\hat{P}) = \left(1 - \frac{n}{N}\right)\frac{S^2}{n}$$

Donde,

$$S^2 = \frac{1}{N-1}\left[\sum_{i=1}^{N} y_i^{\,2} - N\bar{Y}^2\right] = \frac{1}{N-1}\left[\sum_{i=1}^{N} y_i^{\,2} - N\bar{Y}^2\right]$$

$$= \frac{1}{N-1}\left[\sum_{i=1}^{N} Y_i^2 - \left(\sum_{i=1}^{N} Y_i\right) \Big/ N\right] = \frac{NP - NP^2}{N-1} = \frac{NP(1-P)}{N-1}$$

Si N es muy grande, $\frac{N}{N-1} \approx 1$, tenemos que su estimador es,

$$V(\hat{P}) = \left(1 - \frac{n}{N}\right) \cdot \frac{\hat{P}(1-\hat{P})}{n-1}$$

Y su intervalo del $100(1-\alpha)\%$ de confianza es:

$$\hat{P} \pm Z_{(1-\alpha/2)} \sqrt{\left(1 - \frac{n}{N}\right) \cdot \frac{\hat{P}(1-\hat{P})}{n-1}}$$

El total de unidades que tiene la característica A es: $\hat{N}_0 = N\hat{P}$, que es un estimador insesgado, con varianza

$$V(\hat{N}_0) = N^2 V(\hat{P})$$

1.6 Muestreo Sistemático con Arranque Aleatorio

Es la forma de seleccionar la muestra en la cual solamente la primera unidad de la muestra es tomada al azar y el resto se selecciona sistemáticamente. Si N es el número de unidades que constituyen la población es divisible por el de unidades que forman la muestra: $\frac{N}{n} = k$

Podemos considerar la población dividida en n zonas de k unidades y especificar el método de selección de la muestra en la forma siguiente:

- Se selecciona una unidad en la primera zona con probabilidad $\frac{1}{k}$.
- Las n–1 unidades restantes de la muestra son las que ocupan el mismo lugar relativo en las zonas restantes.

Este método tiene grandes ventajas, ya que extiende la muestra a toda la población, recoge el posible efecto de estratificación debido al orden en que figuran las unidades de la población y es de fácil aplicación y comprobación, sin embargo, entre sus inconvenientes figura la posibilidad de aumentar la varianza debido a la periodicidad y el problema teórico que se presenta en la estimación de varianzas, ya que no se puede calcular con una sola muestra.

Cuando las unidades de la población están en un orden aleatorio en el marco, con respecto a los valores de Y_i, el muestreo sistemático es equivalente al muestreo aleatorio simple Y_i, o bien cuando las unidades de la población están ordenadas en el marco en relación a los valores de Y_i, el muestreo sistemático produce varianzas de los estimadores menores que los correspondientes en el muestreo aleatorio simple. Esto se debe a que la muestra queda más dispersa sobre la población. Por otro lado, si las unidades de la población tienen un orden que se refleja en cambios periódicos de los valores de Y_i y el periodo coincide con el valor de *k*, el muestreo sistemático puede producir varianzas mayores a los estimadores, en contraposición al muestreo aleatorio simple. En este caso el problema es que la muestra puede coincidir con todos los valores bajos (o altos) de Y_i, siendo de esta manera poco representativa y con fuertes fluctuaciones de muestra a muestra. Cuando *N* no es divisible entre *k*, el tamaño de muestra es *n* o *n+1* dependiendo de la semilla aleatoria seleccionada, en este caso la media muestral es un estimador sesgado de media poblacional, pero el sesgo es **negligible.**

En este tipo de muestreo no hay forma de estimar $V(\hat{\bar{Y}}_{sis})$, usualmente se utilizan las expresiones del muestreo aleatorio simple, sin embargo cuando se utiliza el muestreo por conglomerados tenemos que:

$$\hat{\bar{Y}}_{sis} = y$$

Con varianza:

$$V(\hat{\bar{Y}}_{sis}) = \frac{\sum_{i=1}^{k}(\bar{Y}_i - \bar{Y})^2}{k} = \frac{k-1}{k}\,\frac{\sum_{i=1}^{k}(\bar{Y}_i - \bar{Y})^2}{k-1} = S_b^2$$

Donde S_b^2 es la varianza entre conglomerados

$$V(\hat{\bar{Y}}_{sis}) = \frac{N-1}{N}S^2 - \frac{k(n-1)}{N}S_w^2$$

donde,

$$S_w^2 = \frac{1}{k(n-1)}\sum_{i=1}^{k}\sum_{j=1}^{n}(Y_{ij} - \bar{Y}_i)^2$$, que es la varianza dentro de conglomerados

Habitualmente se usa el muestreo sistemático para seleccionar unidades secundarias o terciarias, y la varianza de estimadores a ese nivel no se va requerir.

1.7 Muestreo Estratificado

Una muestra es estratificada cuando los elementos de la muestra son proporcionales a su presencia en la población. La presencia de un elemento en un estrato excluye su presencia en otro. Para este tipo de muestreo, se divide a la población en varios grupos o estratos con el fin de dar representatividad a los distintos factores que integran el universo de estudio, donde los estratos se tipifican conforme a su categoría. Este método reside en dividir a la población en *L* subconjuntos llamados estratos con una muestra probabilística en cada uno de ellos de manera independiente de un estrato a otro. Los estratos se forman en función a variables altamente correlacionadas con las variables de estudio, es decir, están constituidos por unidades homogéneas, aunque sean heterogéneas entre estratos, esto implica que cada estrato se muestrea por separado y se obtienen los estimadores de los parámetros para cada uno de los estratos. El efecto de particionar a la población, es decir estratificarla, es reducir la variabilidad de los estimadores, esto es posible si los estratos son muy homogéneos dentro de cada uno de ellos y heterogéneos entre los mismos.

Es conveniente formar estratos, puesto que la disponibilidad de marcos con frecuencia es limitada. De esta forma si para una parte de la población se tiene un buen marco, éste se utiliza para el muestreo y las otras partes de la población se muestrean utilizando otros marcos mas imprecisos y, posiblemente distintos esquemas (diseños) de muestreo. Por otra parte el formar estratos en la población es reducir el costo de localizar la información de las unidades y captación de información semejantes.

En cada estrato es posible utilizar diferentes formas de muestreo

1.7.1 Definición y Notación

En este método de muestreo es válido emplear cualquier estrategia de muestreo probabilístico en cada estrato, incluso puede ser diferentes de un estrato a otro.

N_h	Número de guarderías a seleccionar en el estrato *h*-ésimo; *h*= 1,2,...*L*;
Y_{hi}	Valor de la medición en el elemento *i*-ésimo del estrato *h*-ésimo.

L	Número de estratos.
$N = \sum_{h=1}^{L} N_h$	Total de unidades en la población.
$\bar{Y}_h = \frac{\sum_{i=1}^{N_h} Y_{hi}}{N_h}$	Media poblacional del estrato *h*-ésimo.
$Y_h = N_h \bar{Y}_h = \sum_{i=1}^{N_h} Y_{hi}$	Total poblacional del estrato *h*-ésimo.
$S_h^2 = \frac{\sum_{i=1}^{N_h} (Y_{hi} - \bar{Y})^2}{N_h - 1}$	Varianza poblacional del estrato *h*-ésimo.
$Y = \sum_{h=1}^{L} Y_h = \sum_{h=1}^{L} \sum_{i=1}^{N_h} Y_{hi} = \sum_{h=1}^{L} N_h \bar{Y}_h$	Total de toda la población.

$$\bar{Y} = \frac{Y}{\sum N_h}$$ Media de los valores Y_{hi} en toda la población.

$$W_h = \frac{N_h}{N}, \quad \sum_{h=1}^{L} W_h = 1$$ Proporción del tamaño del estrato *h*-ésimo.

n_h Total de unidades en la muestra del estrato *h*-ésimo.

$$f_h = \frac{n_h}{N_h}$$ Fracción de muestreo en el *h*-ésimo estrato.

$$\bar{y}_h = \frac{\sum_{i=1}^{n_h} y_{hi}}{n_h}$$ Media muestral del estrato *h*-ésimo.

$$s_h^2 = \frac{\sum_{i=1}^{n_h} (y_{hi} - \bar{y}_h)^2}{n_h - 1}$$ Varianza muestral del estrato *h*-ésimo.

1.7.2 Estimadores y Varianzas

La media poblacional total de $\bar{Y}$ está dada por la suma ponderada de la media de todos los estratos:

$$\bar{Y} = \frac{\sum_{h=1}^{L}\sum_{i=1}^{N_h} y_{hi}}{N} = \sum_{h=1}^{L} \frac{1}{N} N_h \sum_{i=1}^{N_h} \frac{y_{hi}}{N_h} = \sum_{h=1}^{L} \frac{N_h}{N} \bar{Y}_h = \sum_{h=1}^{L} W_h \bar{Y}_h$$

Para obtener el estimador de $\bar{Y}$ se considera $\bar{y}_h$ como estimador de $\bar{Y}_h$, sustituyendo en la expresión anterior, tenemos que:

$$\hat{\bar{Y}} = \sum_{h=1}^{L} W_h \bar{y}_h$$

Si en cada uno L de los estratos $\bar{y}_h$ es estimador insesgado de $\bar{Y}_h$ entonces $\bar{y}$ es un estimador insesgado de $\bar{Y}$.

$$E(\bar{y}) = E\sum_{h=1}^{L} \frac{N_h}{N} \bar{y}_h = \sum_{h=1}^{L} \frac{N_h}{N} E(\bar{y}_h)$$
$$= \sum_{h=1}^{L} \frac{N_h}{N} \bar{Y}_h = \sum_{h=1}^{L} W_h \bar{Y}_h = \bar{Y}$$

En el muestreo estratificado, suponiendo $\bar{y}_h$ insesgado para cada estrato las muestras se obtienen de manera independiente, la varianza de $\bar{y}$ como estimador de $\bar{Y}$,

$$V(\bar{y}) = V\left(\sum_{h=1}^{L} \frac{N_h}{N}\bar{y}_h\right) = \sum_{h=1}^{L} \frac{N_h^2}{N^2} V(\bar{y}_h) = \sum_{h=1}^{L} W_h^2 V(\bar{y}_h)$$

Utilizando muestreo aleatorio simple en todos los estratos tenemos que la varianza del estimador de la media $\bar{y}_h$ tiene la expresión siguiente:

$$\begin{aligned} V(\bar{y}) &= \sum_{h=1}^{L} W_h^2 V(\bar{y}_h) = \sum_{h=1}^{L} W_h^2 \left(1 - \frac{n_h}{N_h}\right)\frac{S_h^2}{n_h} \\ &= \sum_{h=1}^{L} \frac{N_h^2}{n_h^2}\left(1 - \frac{n_h}{N_h}\right)\frac{S_h^2}{n_h} \\ &= \frac{1}{N^2}\sum_{h=1}^{L}\frac{N_h^2 S_h^2}{n_h} - \frac{1}{N^2}\sum_{h=1}^{L} N_h S_h^2 \end{aligned}$$

Si se ignora el factor de corrección por finitud $\left(1 - \frac{n_h}{N_h}\right)$ la expresión se simplifica en:

$$V(\bar{y}) = \frac{1}{N^2}\sum_{h=1}^{L}\frac{N_h^2 S_h^2}{n_h}$$

El estimador del total Y se obtiene mediante la multiplicación del estimador de la media total por el tamaño de la población: $\hat{Y} = N\bar{y}$

La varianza del estimador del total se calcula mediante el producto del cuadrado del tamaño de la población total por la varianza de la media: $V(\hat{Y}) = N^2 V(\bar{y})$

Para asignar el tamaño de muestra correspondiente a cada estrato, se divide el tamaño total de la muestra entre los L estratos. De este modo la expresión n_h sería la siguiente: $n_h = \frac{n}{L}$.

Al sustituir $n_h = \frac{n}{L}$ en la fórmula general, se obtiene forma particular:

$$V(\bar{y}) = \frac{L}{N^2}\sum_{h=1}^{L}\frac{N_h^2 S_h^2}{n} - \frac{1}{N^2}\sum_{h=1}^{L} N_h S_h^2 .$$

De la misma manera, para asignar el tamaño de muestra proporcional al tamaño del estrato, se parte de la relación de proporcionalidad que iguala la razón del tamaño del estrato respecto al tamaño de la población con la razón del tamaño de muestra en el estrato respecto al tamaño total de la muestra $\frac{n_h}{n} = \frac{N_h}{N}$, de aquí se despeja n_h para obtener la fórmula de afijación, es decir, $n_h = \frac{N_h}{N} n$. Así se obtendrá la expresión correspondiente a la varianza de $\bar{y}$, considerando la afijación proporcional.

Primero se sustituye $n_h = \frac{N_h}{N} n$ en la expresión de varianza general de la siguiente manera:

$$V(\bar{y}) = \frac{1}{N^2} \sum_{h=1}^{L} \frac{N_h^2 S_h^2}{\frac{N_h}{N} n} - \frac{1}{N^2} \sum_{h=1}^{L} N_h S_h^2$$

$$\Rightarrow \qquad = \frac{1}{N} \sum_{h=1}^{L} \frac{N_h S_h^2}{n} - \frac{1}{N^2} \sum_{h=1}^{L} N_h S_h^2$$

1.7.3 Determinación del Tamaño de Muestra

Para obtener el tamaño de muestra *n*, se parte con la relación de precisión, confianza y varianza: $d^2 = Z^2_{(1-\alpha/2)} V(\bar{y})$, se despeja la $V(\bar{y})$ y se le asignan valores a la precisión y el coeficiente de confianza $Z^2_{(1-\alpha/2)}$. La varianza quedará igualada a una constante que llamaremos D^2 y así se obtiene la varianza deseada:

$$\Rightarrow \qquad V(\bar{y}) = \frac{d^2}{Z^2_{(1-\alpha/2)}} = D^2$$

$$\Rightarrow \qquad V(\bar{y}) = D^2$$

Escribiendo de forma explícita $V(y)$, se despeja *n* para obtener el tamaño de muestra.

$$D^2 = \frac{L}{N^2} \sum_{h=1}^{L} \frac{N_h^2 S_h^2}{n} - \frac{1}{N^2} \sum_{h=1}^{L} N_h S_h^2$$

$$\Rightarrow \qquad N^2 D^2 + \sum_{h=1}^{L} N_h S_h^2 = \frac{L}{n} \sum_{h=1}^{L} N_h^2 S_h^2$$

$$\Rightarrow \qquad n = \frac{L \sum_{h=1}^{L} N_h^2 S_h^2}{N^2 D^2 + \sum_{h=1}^{L} N_h S_h^2}$$

El tamaño de muestra total n, suponiendo la varianza deseada D^2, se obtiene despejando n de la formula de varianza $\bar{y}$ según el criterio de afijación proporcional

$$D^2 = \frac{1}{N}\sum_{h=1}^{L}\frac{N_h S_h^2}{n} - \frac{1}{N^2}\sum_{h=1}^{L} N_h S_h^2$$

$$\Rightarrow \qquad N^2 D^2 + \sum_{h=1}^{L} N_h S_h^2 = N\sum_{h=1}^{L}\frac{N_h S_h^2}{n}$$

$$\Rightarrow \qquad n = \frac{N\sum_{h=1}^{L} N_h S_h^2}{N^2 D^2 + \sum_{h=1}^{L} N_h S_h^2}$$

CAPITULO II. METODOLOGÍA DE MUESTREO APLICADA AL ANÁLISIS DE LA OFERTA Y DEMANDA DEL SEGURO DE GUARDERÍAS.

2.1 Referencias metodológicas

Para cumplir con los objetivos y las metas del proyecto, a continuación se muestran las referencias metodológicas que se plantearon.

- **Unidad de análisis:** Los niños menores de 5 años que sean residentes de la Guarderías seleccionadas en la muestra.
- **Unidad de observación:** Las Guarderías seleccionadas por modalidad y sus salas de atención.
- **Método de recolección:** Captar información por medio de un cuestionario general mediante entrevista directa.

2.2 Selección del Método de Muestreo

El procedimiento utilizado es el de encuesta por muestreo estratificado, donde las unidades del marco son de inclusión obligatoria (estrato forzoso). La unidad estadística de muestreo es la guardería, también considerada como la unidad informante, ya que al estar perfectamente definida y localizada y disponer de los datos contables y de la capacidad instalada e inscripción de los niños con derecho a esta prestación, se facilita la respuesta y se obtiene información homogénea. Cada unidad de muestreo se realizó de forma independiente para cada estrato, el procedimiento de selección se aplicó de acuerdo a la modalidad, donde cada uno de los estratos se seleccionó con probabilidad proporcional al tamaño.

Dentro de cada estrato se seleccionó una muestra probabilística de las guarderías y luego en cada una de las modalidades se escogió, mediante sorteo aleatorio simple, cuáles serían las guarderías que integrarían la muestra. Cabe recalcar que la selección de la muestra se realizó de manera independiente por entidad, dominio y estrato y que el procedimiento de selección varió de acuerdo con el dominio.

Las guarderías fueron elegidas mediante selección proporcional a su tamaño (en base a la inscripción de los infantes).

El sorteo hace que la relación de las guarderías sea representativa de su participación en el universo (a nivel del Distrito Federal), pues aquellas guarderías con más inscripción de infantes tendrían más probabilidad de ser seleccionadas que aquellas más pequeñas. Esto obliga a respetar la relación de las guarderías seleccionadas.

2.3 Criterios del Esquema de Muestreo.

El esquema de muestreo para el estudio de la oferta y la demanda de las guarderías es probabilístico y estratificado.

- **Probabilístico.** Porque todas las unidades de muestreo tienen una probabilidad conocida y distinta de cero de ser seleccionadas, es decir, se escogió de cada modalidad cierto número de guarderías que represente a la población a inferir.
- **Estratificado.** Porque las unidades de muestreo con características similares se agrupan para formar estratos, estos están formados por las distintas modalidades que maneja el IMSS.

2.3.1 Prueba Piloto

Se realizó una prueba piloto para estimar la varianza, la cual era necesaria para calcular el tamaño de muestra adecuado.

Se tomó n=10 fija para cada de región y a su vez se escogieron aleatoriamente las guarderías para poder estimar la varianza de los niños inscritos.

2.3.2 Marco Muestral

La encuesta se aplicará en las guarderías seleccionadas mediante una muestra aleatoria, extraída de un marco muestral que representa a la población bajo estudio, para su conformación y estructura, se utilizó un diseño probabilístico, estratificado.

El marco de la encuesta que se utiliza para la oferta del seguro de guarderías esta conformado por el directorio nacional que proporciona el Instituto Mexicano del Seguro Social que contiene información sobre los servicios de la población Infantil, lo que permite su estratificación por modalidad. También figuran en ese directorio datos sobre la identificación y localización de las unidades estadísticas, que son necesarios para un correcto relevamiento de la información.

Una vez estratificado el marco, se procedió a la selección de las guarderías mediante un muestreo aleatorio con intervalo constante.

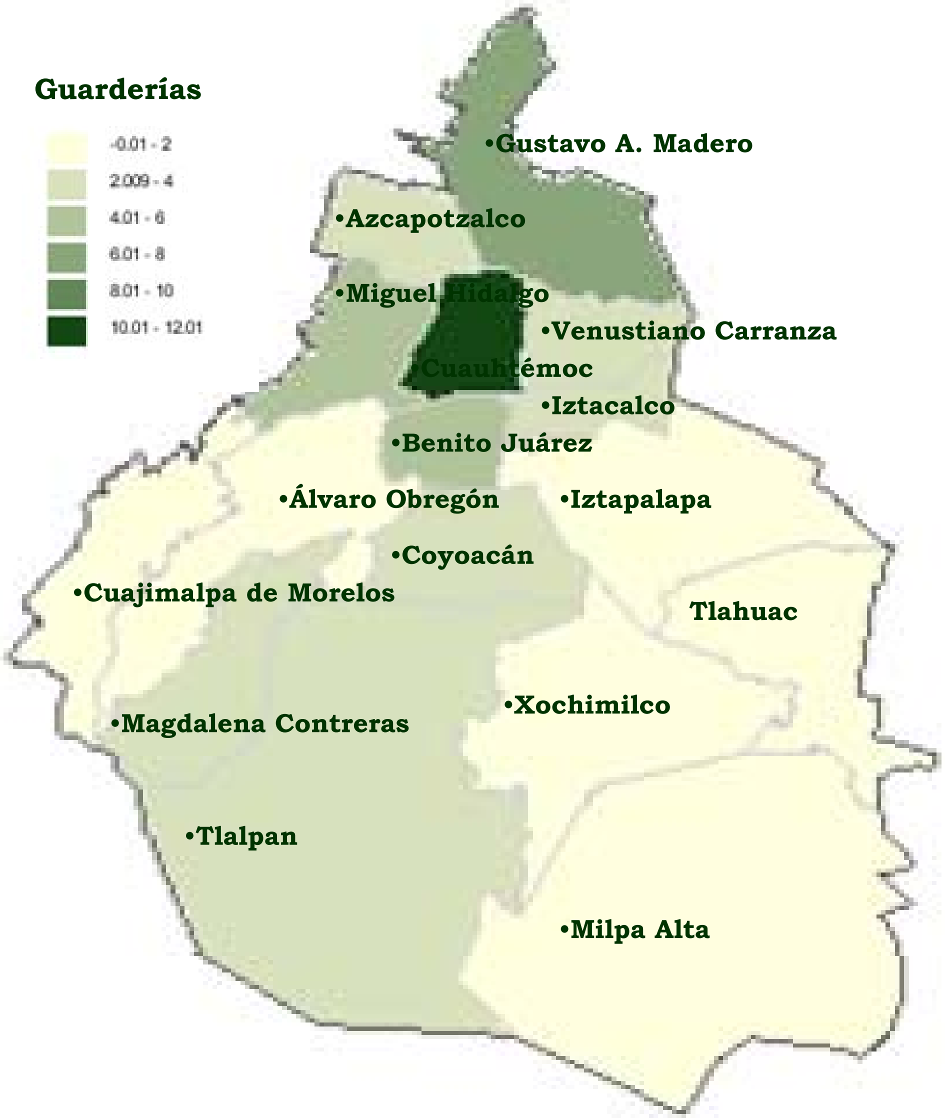
2.3.3 Diseño Muestral

El diseño de la muestra para el análisis de la oferta y la demanda del seguro de guarderías se caracteriza por ser probabilístico, en consecuencia, los resultados obtenidos se generalizan a toda la población de niños de edades de 43 días a 4 años que sean derechohabientes del Instituto Mexicano del Seguro Social y tengan derecho a las mismas. También es posible medir los errores de las estimaciones obtenidas de la encuesta. A su vez, el diseño es estratificado. Las guarderías han sido estratificadas en base a la modalidad que maneja el IMSS, por facilidad de ubicación de las guarderías y la oferta que ofrecen.

2.3.4 Estratificación de las Unidades de Muestreo

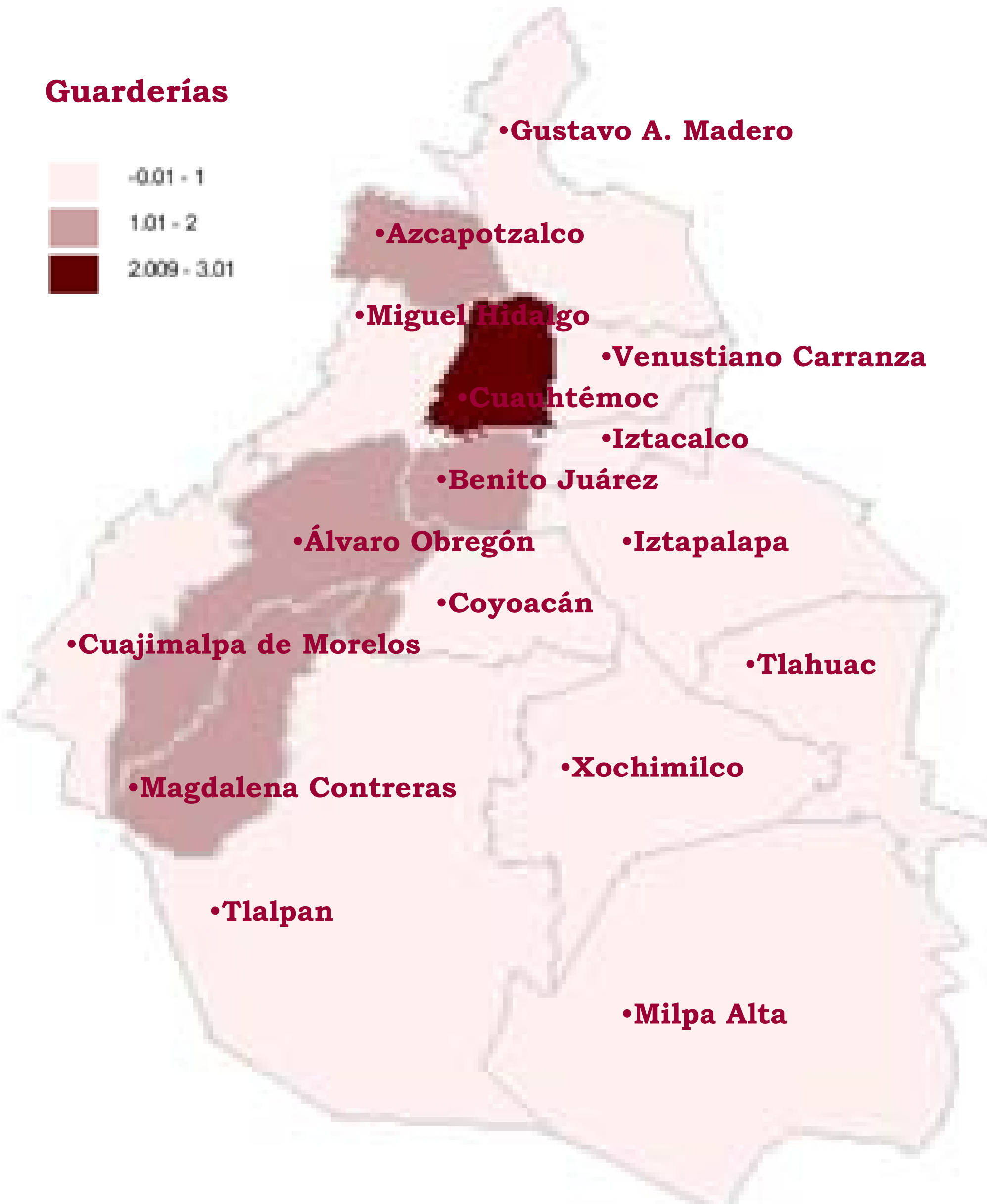
- Ordinarias
- Madres IMSS
- Esquema Único

Gráfica 1.1
Mapeo de guarderías de modalidad ordinaria en el Distrito Federal por demarcación política



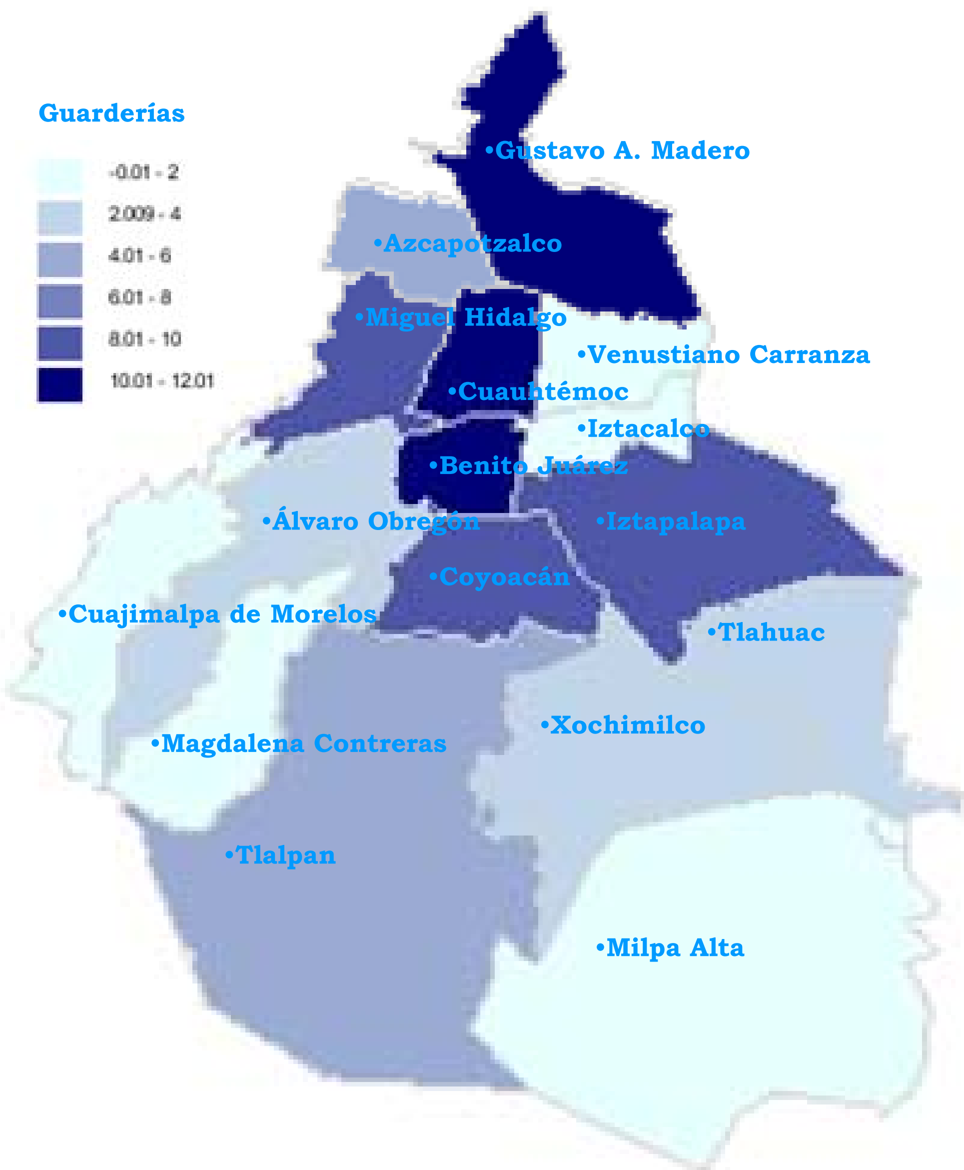
Fuente: Elaboración propia a partir de información estadística del IMSS.

Gráfica 1.2
Mapeo de guarderías de modalidad Madres IMSS en el Distrito Federal por demarcación política



Fuente: Elaboración propia a partir de información estadística del IMSS.

Gráfica 1.3
Mapeo de guarderías de modalidad de esquema único en el Distrito Federal por demarcación política



Fuente: Elaboración propia a partir de información estadística del IMSS.

2.3.5 Cálculo del tamaño de la muestra

El tamaño de la muestra permite dimensionar la participación y distribución de las unidades de análisis y la población objetivo, de esta forma un tamaño muestral adecuado permite un buen acercamiento a los fenómenos estudiados y el análisis de los resultados permite proporcionar información desglosada a nuestros usuarios para el Distrito Federal en sus diferentes ámbitos. El tamaño de muestra se calculó tomando como variable de referencia a la proporción de los niños con derecho a esta prestación, la expresión para calcular el tamaño de muestra fue la siguiente:

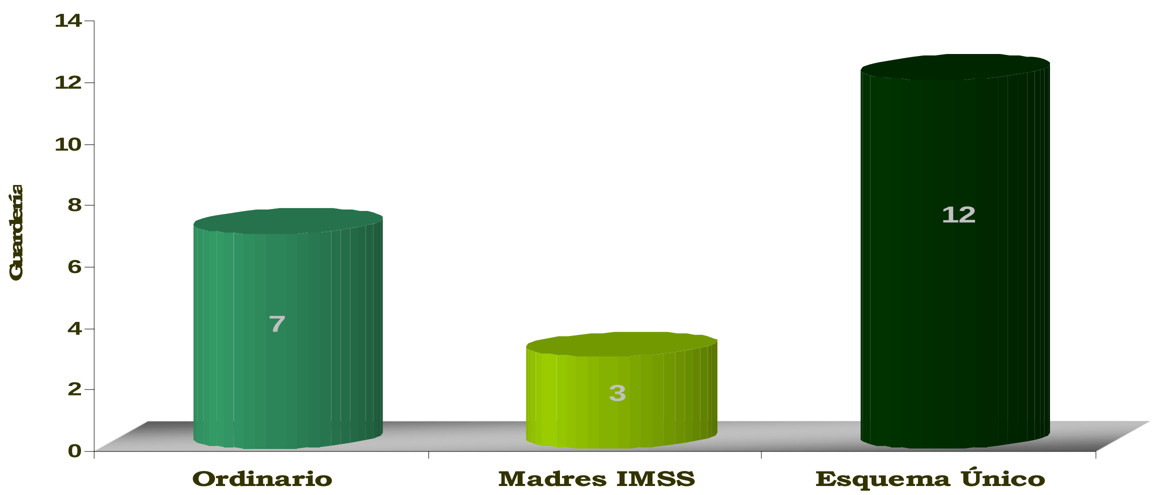
$$n = \frac{N \sum_{h=1}^{L} N_h S_h^2}{N^2 D^2 + \sum_{h=1}^{L} N_h S_h^2}$$

Considerando una confianza del 95% y un error relativo no mayor al 1%.

Estas condiciones dieron por resultado un tamaño de muestra de 22 unidades de muestreo (guarderías) de las 144 que existen en el Distrito Federal.

A continuación se presenta la gráfica de la distribución de la muestra en guarderías por modalidad y el listado de guarderías seleccionadas en la muestra.

Gráfica 1.4
Distribución de la muestra por Modalidad de Guarderías IMSS



Fuente: Estimaciones por muestreo, 2007.

Tabla 1.1
Listado de Guarderías Seleccionadas en la Muestra por Modalidad y Delegación

Modalidad de la Guardería	Delegación
Ordinario 0001	Miguel Hidalgo
Ordinario 0004	Cuauhtémoc
Ordinario 0010	Azcapotzalco
Ordinario 0021	Venustiano Carranza
Ordinario 0025	Tlalpan
Ordinario 0031	Benito Juárez
Ordinario 0034	Cuauhtémoc
Madres IMSS IV	Azcapotzalco
Madres IMSS VI	Magdalena Contreras
Madres IMSS VIII	Cuauhtémoc
U - 0046 Colegio Pedagogíco de la Ciudad de México, S.C.	Gustavo A. Madero
U - 0047 Colegio de las Baleares S.C.	Miguel Hidalgo
U - 0049 Colegio Pedagógico de la Cuidad de México, S.C.	Xochimilco
U - 0055 Estancia Infantil Tercer Milenio, S.C.	Cuauhtémoc
U - 0059 Instituto de las Baleares, S.C.	Gustavo A. Madero
U - 0182 Servicios Integrales Proinfancia, S.C.	Benito Juárez
U - 0287 Instituto Angeli, S.C.	Benito Juárez
U - 0538 Centro de Desarrollo Infantil Izkalotl, S.C.	Benito Juárez
U - 0541 Dreams House, S.C.	Coyocán
U - 0618 Guard Centro Histórico de la Ciudad de México,D.F.	Tlahuác
U - 0629 Proinfancia Teresita, S.C.	Alvaro Obregon
U - 0698 Cto.de Desarrollo Infantil Comunitario Yoltzin,S.C	Benito Juárez

Fuente: Recopilación de datos del directorio del IMSS a partir de cálculos de muestra.

2.3.6 Variables

Las variables elegidas e incluidas en el marco conceptual para este estudio, se agrupan de acuerdo con los temas considerados para la captación de datos y son las que proporcionan la información requerida para analizar la oferta y la demanda del seguro de guarderías, a partir de lo cual se valoraron los siguientes aspectos[11]:

[11]Véase cuestionario en anexo

- Edad de los infantes que comprenden entre los 43 días de nacidos a 4 años de edad con derecho a la prestación, donde la Institución los cataloga por sala de atención.
- Edad quinquenal de los derechohabientes que utilizan el servició.
- Sexo: Masculino y Femenino.
- Capacidad Máxima: Número de menores que se pueden atender.
- Niños atendidos en cada instancia.
- Niños en trámite, es decir menores que estén en espera para se atendidos.
- Lugar de residencia del tutor del menor.
- Costo del servicio.
- Domicilio de la Guardería a encuestar.
- Modalidad de la Guardería.

2.4 Población Objetivo.

- Objetivo: Generar información Estadística para dimensionar y caracterizar las demandas de los usuarios de esta prestación en el Distrito Federal y de manera general, la atención que se da a los usuarios.
- Población Objetivo: La población objeto de estudio son los menores de edad entre 43 días y 4 años que ocupan las estancias del IMSS ubicadas dentro del Distrito Federal, para percibir y analizar la atención brindada por el Instituto (oferta) y la espera de niños para ser atendidos (demanda).

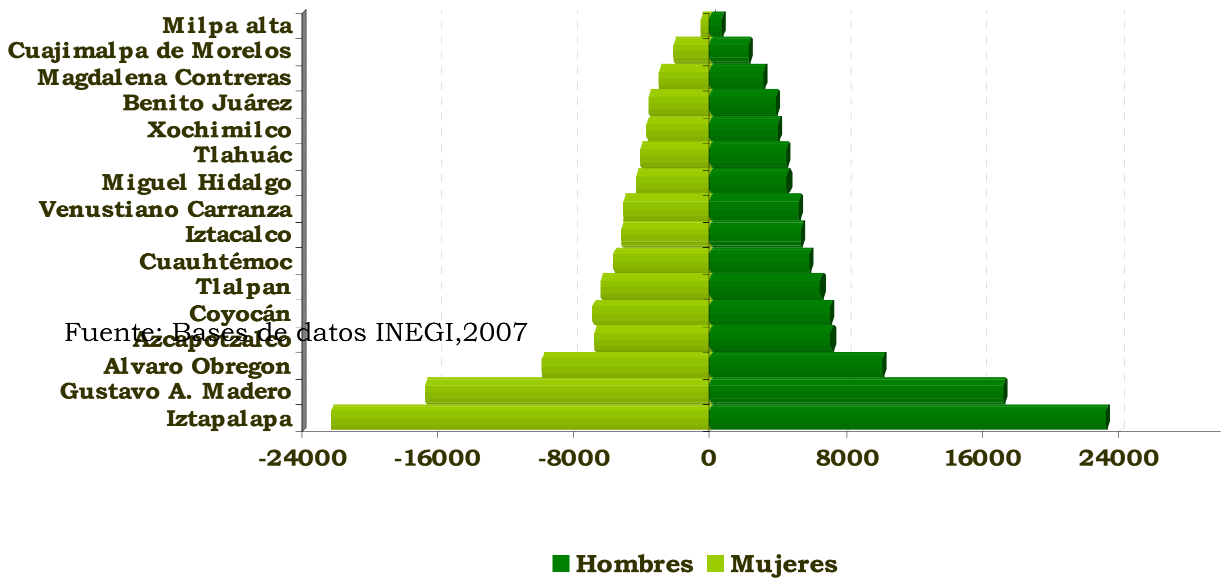
2.4.1 Características de la Población

Los niños que tienen derecho a la prestación de las instalaciones del IMSS, esta conformada por las madres trabajadoras aseguradas, los padres viudos o divorciados con la custodia legal de su hijo, o cualquier otro trabajador que ejerza la patria potestad y la custodia de un menor, durante su jornada laboral, cuidado infantil de calidad a sus hijos pequeños de 43 días a cuatro años de edad, sanos o con alguna discapacidad leve[12] o moderada[13], además de no padecer alguna enfermedad contagiosa o crónica y que tienen derecho al servicio de guarderías.

[12] Una **discapacidad leve** se presenta cuando la reducción de la capacidad del menor para desempeñar sus actividades cotidianas es mínima y no interfiere en su rendimiento. Estos niños, en general, no requieren de una atención especializada.

[13] Una **discapacidad moderada** se presenta cuando la reducción de la capacidad del menor limita parcialmente sus actividades y productividad. Estos niños requieren un manejo y atención especializada que les permita adecuarse al medio y compensar sus limitaciones para lograr su integración plena a la sociedad

Gráfica 1.5
Población de niños de 0 a 4 años de edad con derecho a la prestación de Guarderías del IMSS por delegación



Fuente: Bases de datos INEGI,2007

2.4.2 Por Modalidad

Se consideraron las distintas modalidades que maneja el IMSS: Ordinarias, Madres IMSS y Esquema Único En el siguiente tabulado se presenta la distribución de las guarderías que están en el Distrito Federal, por cada Delegación, según su modalidad.

Tabla 1.2
Distribución de Guarderías por delegación, según modalidad

Delegación	Ordinarias	Madres IMSS	Esquema Único
Distrito Federal	44	7	96
Alvaro Obregon	0	1	3
Azcapotzalco	2	1	6
Benito Juárez	4	1	11
Coyocán	3	0	8
Cuajimalpa de Morelos	0	0	1
Cuauhtémoc	12	3	14
Gustavo A. Madero	6	0	16
Iztacalco	3	0	1
Iztapalapa	1	0	10
Magdalena Contreras	2	1	2
Miguel Hidalgo	5	0	10
Milpa alta	0	0	0
Tlahuác	0	0	3
Tlalpan	3	0	5
Venustiano Carranza	3	0	3
Xochimilco	0	0	3

Fuente: IMSS, Guarderías

Tabla 1.3
Distribución de Guarderías seleccionadas en la muestra por delegación, según modalidad

Delegación	Ordinarias	Madres IMSS	Esquema Único
Total	**7**	**3**	**12**
Alvaro Obregon	0	0	1
Azcapotzalco	1	1	0
Benito Juárez	1	0	4
Coyocán	0	0	1
Cuajimalpa de Morelos	0	0	0
Cuauhtémoc	2	1	1
Gustavo A. Madero	0	0	2
Iztacalco	0	0	0
Iztapalapa	0	0	0
Magdalena Contreras	0	1	0
Miguel Hidalgo	1	0	1
Milpa alta	0	0	0
Tlahuác	0	0	1
Tlalpan	1	0	0
Venustiano Carranza	1	0	0
Xochimilco	0	0	1

Fuente: Estimaciones de cálculo de muestra a partir del directorio de guarderías del IMSS

2.4.3 Por grupo de edad según IMSS

Los niños(as) son ubicados en salas de atención según su edad y nivel de desarrollo, adecuando las actividades a las capacidades y habilidades de cada nivel.

Tabla 1.4
Salas de atención por edades y nivel de desarrollo

Sala de atención	Edad de los niños
Lactantes "A"	43 días a 6 meses
Lactantes "B"	7 meses a 12 meses
Lactantes "C"	13 meses a 18 meses
Maternales "A"	19 meses a 24 meses
Maternales "B1"	25 meses a 30 meses
Maternales "B2"	31 meses a 36 meses
Maternales "C1"	37 meses a 42 meses
Maternales "C2"	43 meses a 48 meses

Fuente: Dirección de Prestaciones Económicas y Sociales, IMSS.

2.5 Periodo de Estudio

Período de referencia: La recolección de información individual comprende, en su conjunto, un mes natural entre octubre y noviembre de 2007, natural por considerarse más adecuado para el estudio de las variables objetivo del proyecto.

El período de referencia de los resultados es conciso. Ello es consecuencia de que la mayor parte de las variables van referidas a períodos delimitados en función de las características de las distintas modalidades analizadas.

2.6 Panorama General de los Servicios para el Cuidado de la Población Infantil

La guardería no solamente ayuda a resolver el aspecto social de la incorporación de la mujer al trabajo sino que es una aportación muy valiosa para garantizar que los niños están seguros, con una nutrición apropiada, estimulados correctamente desde el punto de vista de su desarrollo y rodeados de cariño, favoreciéndose así su proceso de socialización. El Instituto Mexicano del Seguro Social brinda los servicios de guardería a las madres trabajadoras aseguradas, ofreciendo un espacio educativo-formativo para sus hijos en la primera infancia, mientras ellas desarrollan sus labores durante la jornada de trabajo.

Los servicios de guardería del sistema IMSS han ido ampliándose y perfeccionándose con el objeto de proporcionar más y mejores guarderías para el cada vez mayor número de mujeres que se incorporan a los mercados laborales.

El IMSS estipula que las niñas y niños desde los 43 días hasta los 4 años de edad tienen derecho al servicio de guarderías, siempre y cuando se cumplan con los requisitos siguientes:

- Ser madres trabajadoras.
- Ser padres viudos o divorciados que por resolución judicial tengan la custodia de sus hijos.
- Cualquier asegurado que ejerza la patria potestad o la custodia de un menor.

El IMSS proporciona los servicios de guardería ya sea en sus propias instalaciones o bien a través de prestadores de servicio.

Las guarderías del IMSS incluyen los servicios de aseo, alimentación, cuidado de la salud, educación y recreación de las menores. La Ley del Seguro Social define que el financiamiento a la prestación en especie del servicio de guarderías se realice con por lo menos 80 por ciento de los ingresos de las aportaciones al Seguro de Guarderías y Prestaciones Sociales. En sus guarderías se cuida y se fortalece la salud del niño y su buen desarrollo futuro, la formación de sentimientos de adhesión familiar y social, la adquisición de conocimientos que promuevan la comprensión, el empleo de la razón y de la imaginación; el respeto a la diversidad de habilidades, antecedentes sociales y culturales de todos y cada uno de los niños y niñas que asisten a sus guarderías y la construcción de hábitos higiénicos y de sana convivencia y cooperación, conforme a su edad y realidad social.

El IMSS ha diseñado todo un concepto que permite que todas las guarderías tengan una estructura uniforme.

- Ha establecido los perfiles que debe tener el personal que labora en la guardería y atiende cada uno de los servicios, haciendo especial énfasis en la vocación de servicio y calidez hacia los niños.
- Determina quiénes y cuántos deben atender cada servicio.
- Ha establecido los programas que le permitan al personal conocer el qué hacer y cómo hacer con cada una de las actividades que se realizan en la guardería.
- Tiene un programa permanente de asesoría a las guarderías y supervisa que cumplan con los programas y objetivos establecidos.

2.7 El Sistema de Guarderías IMSS

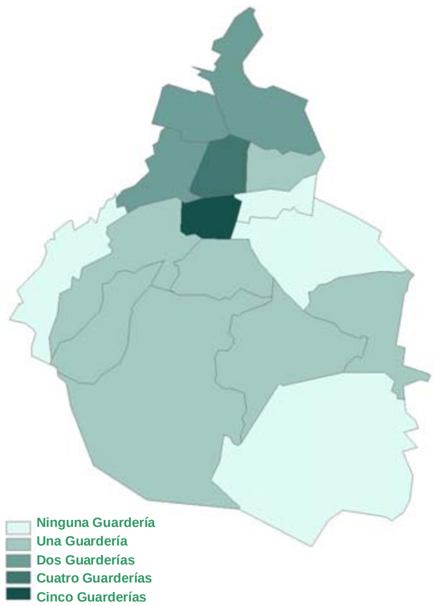
Uno de los objetivos prioritarios para el Instituto ha sido extender el servicio de guarderías y en 1997 se comenzó a adoptar un esquema subrogado de prestación del servicio llamado vecinal comunitario, cuyos costos son sustancialmente inferiores a los que registran los otros esquemas, no requiere inversión de capital, promueve la eficiencia y aumenta la flexibilidad del servicio. Anteriormente el servicio de guarderías se prestaba a través de cinco modalidades: Ordinaria, Madres IMSS, Participativas, Vecinal Comunitaria y del Campo, las que tienen importantes diferencias en costos. A este respecto la política presupuestal consiste en reconocer dicho diferencial de costos, pero procurando disminuir la variación de costos dentro de una misma modalidad. En el año 1999 se identificaron problemas derivados de la heterogeneidad en las tarifas pagadas por los servicios de guardería subrogados a pesar de que los estándares de servicio requeridos eran los mismos para todos los casos. Dada esta situación, además de la inequidad que representaba para los proveedores, percibieron que podría derivar en problemas para la evolución y estabilidad del sistema, por lo cual se adoptó una política de homologación entre las modalidades Participativas, Vecinal Comunitaria y del Campo a una sola.

2.8 Análisis de la Cobertura Geográfica

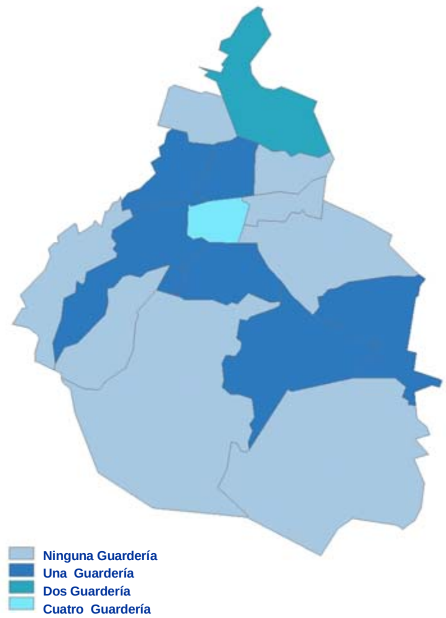
En esta encuesta se generará información con cobertura geográfica del Distrito Federal y permitirá contar con representatividad a nivel Distrito Federal, como se presenta en los siguientes mapas.

Gráfica 1.6
Cobertura Geográfica

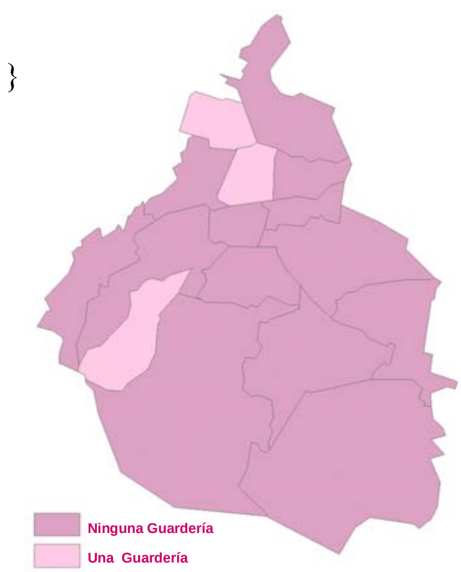
Cobertura Geográfica de guarderías del IMSS seleccionadas en la muestra de todas las Modalidades



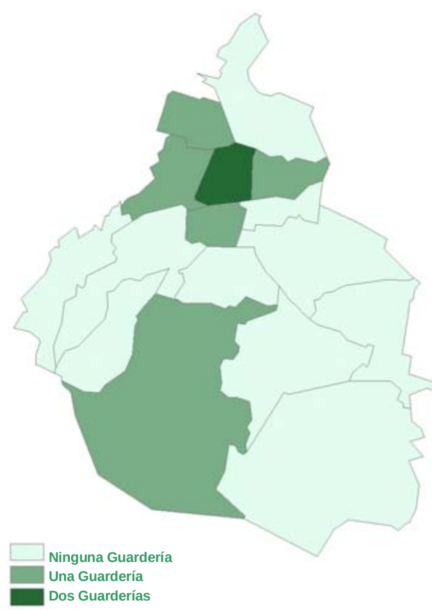
Cobertura Geográfica de guarderías del IMSS seleccionadas en la muestra de modalidad Ordinarias



Cobertura Geográfica de guarderías del IMSS seleccionadas en la muestra de modalidad Madres IMSS



Cobertura Geográfica de guarderías del IMSS seleccionadas en la muestra de modalidad Ordinaria



Fuente: Recopilación de datos del directorio del IMSS a partir de cálculos de muestra.

2.9 Modalidades de Atención

El Sistema de Guarderías del IMSS otorga sus servicios mediante tres modalidades de atención: i) prestación directa; ii) prestación indirecta; y, iii) esquemas de subrogación de servicios con reversión de cuotas.

El primero de estos esquemas de atención se brinda a través de 44 guarderías Ordinarias[14] y 7 guarderías Madres IMSS[15]. En el segundo de estos esquemas el Instituto celebra convenios con prestadores de servicios que poseen guarderías con características inmobiliarias, legales y de funcionamiento operativo que cumplen con los estándares del Instituto; actualmente se encuentran en esta modalidad 96 guarderías en Esquema Único[16] y una de Esquema Integradoras[17]. En lo que se refiere a los esquemas de subrogación de servicios con reversión de cuotas, el Instituto tiene la posibilidad de celebrar convenios de éste tipo con patrones que tienen instaladas guarderías en sus empresas o establecimientos, las cuales deben cumplir con los requisitos establecidos en las normas institucionales para tal efecto.

[14]Ordinarias: Son aquellas que se ubican dentro de las Instalaciones del Seguro Social con mayor capacidad

[15]Madres IMSS: Exclusivas para madres trabajadoras.

[16] Esquema Único: Contemplan el esquema Participativo, Vecinal Comunitario y de Campo, operan como Sociedad Civil a través de una concesión del IMSS

[17] Atención a niños con discapacidad.

CAPITULO III. ÁNALISIS DE RESULTADOS E INFORME DE INVESTIGACIÓN

3.1 Descripción de la Demanda Potencial de Niños

De acuerdo a los datos reportados por el II Conteo de Población y Vivienda 2005 la demanda potencial de niños con derecho a la prestación de guarderías es de alrededor de 2,400,000. Tal y como se esperaba en la hipótesis planteada para esta investigación. Sin embargo, al realizar el presente estudio se observan diversas situaciones inesperadas referentes a la relación oferta-demanda, que se mencionan a continuación:

> 1ª.- La demanda potencial de niños no guarda una relación proporcional con la demanda real, esto implica que a pesar de que la capacidad instalada no es suficiente respecto a la demanda potencial, la demanda real representa el 20% de la oferta.
>
> 2a.- El resultado anterior permite reflexionar sobre los factores externos que influyen notablemente en la determinación de los padres para hacer uso de esta prestación.
>
> 3a.- El nivel socioeconómico de los derechohabientes es inversamente proporcional a la demanda, afectando en la utilización del servicio, dado que los derechohabientes con menos recursos económicos presentan mayor ausentismo; sin embargo, aquellos con mayores recursos, no utilizan el servicio, pues disponen de sus recursos para incorporar a sus hijos en servicios particulares.
>
> 4a.- La observación anterior se refleja en ausentismo, que a su vez representa un costo para la institución y para las guarderías, dado que se tiene en cuenta el volumen regular de utilización, lo que implica contar con recursos suficientes para su atención, por lo que al registrarse inasistencias, el impacto se sufre en la dotación de bienes para el desarrollo de las actividades cotidianas de las guarderías, reduciéndose su disponibilidad.

En el Distrito Federal la concentración geográfica de la población de acuerdo a su ocupación y nivel socioeconómico determinará también el comportamiento de la demanda de acuerdo a la ubicación de las guarderías. Por ejemplo, en aquellos lugares en donde los comercios y las empresas tienen mayor presencia, se espera un mayor número de derechohabientes que además de tener derecho a la prestación, requieren efectivamente hacer uso de ella.

Con referencia a la los derechohabientes, cabe mencionar los siguientes puntos:

- La inseguridad se ha convertido en un factor determinante en la disminución de la demanda;

- Los medios de comunicación en reiteradas ocasiones han manifestado malos tratos a los infantes y el mal uso de las instalaciones, en consecuencia los padres con las posibilidades económicas preferirán hacer uso de instituciones privadas a pesar de contar con la prestación.

- Relacionado directamente con la inseguridad, el género se ha convertido en otro factor determinante en la demanda del servicio. Sobre esta afirmación se muestra que los padres con hijos varones delegarán con mucha mayor facilidad su cuidado a las guarderías, mientras que para las mujeres se prefiere encargar el cuidado a familiares o instituciones privadas.

3.2 Características generales

En el mercado laboral mexicano se ha incrementado la participación y permanencia de las mujeres en los mercados laborales remunerados, que representa una tasa de participación económica cercana al 40 por ciento por tanto se ven obligadas a plantearse decisiones que en muchos casos adquieren un carácter imperioso debido a los conflictos y negociaciones que se generan al interior de la familia, los cuales adquieren una dimensión social más amplia cuando se manifiestan otros problemas como la desatención de los hijos, las enfermedades, la desnutrición, la falta de concentración y el estrés en el trabajo[18].

Algunos padres de familia han encontrado diversas maneras que les han permitido combinar el cuidado de sus hijos e hijas con el trabajo, en muchas ocasiones formando y fortaleciendo redes sociales con el apoyo de otras mujeres dentro y fuera de la misma familia, y en otras tantas, compartiendo las responsabilidades familiares con los padres[19].

No obstante, no todas las madres y los padres cuentan con estas redes, es por ello que son necesarias las guarderías como parte del sistema de protección social que el Estado construye para el fomento del trabajo. Según

[18] Brachet, V. (1996) “Trabajo materno y salud infantil: una guía teórica para las políticas sociales”, pp. 59-86 en *El papel del trabajo materno en la salud infantil* (coord.) Claudio Stern. México DF: Population Council/El Colegio de México

[19] Medina Macías, 2005

la Organización Internacional del Trabajo, los servicios de guardería que ayudan a los padres que trabajan, particularmente a las mujeres, permite conciliar sus responsabilidades parentales y familiares con un empleo remunerado y se constituye en un factor esencial para la consecución de la igualdad de género y la equidad social.

La guardería no solamente ayuda a resolver el aspecto social de la incorporación de la mujer al trabajo, sino que es una aportación muy valiosa para el desarrollo de los niños al permitir su socialización, estimular su psicomotricidad y aprendizaje y garantizar que están seguros y con una nutrición apropiada. Para los padres, las ventajas radican en que disminuyen el riesgo de caer en la pobreza, evitan la pérdida de capital humano, fomentan su permanencia en el mercado laboral y mejoran el aprovechamiento de sus capacidades.

Por estas razones, en el Distrito Federal, el servicio de guarderías es una prestación social que se otorga a la mujer trabajadora y al trabajador solo, durante la jornada laboral, para la atención y cuidado de los hijos en la primera infancia. Su reconocimiento como prestación social se encuentra en los artículos 201 al 207 y 211 al 213 de la Ley del Seguro Social capítulo VII[20].

El principal proveedor de este servicio, es el Instituto Mexicano del Seguro Social (IMSS). No se cuenta exactamente con el número de derechohabientes que utilizan las guarderías, debido a que depende del patrón otorgar dicha prestación, ese, es el fundamento de este estudio.

Al mes de diciembre del 2007 el Sistema de Guarderías IMSS en el Distrito Federal habla 24,323 niños y niñas inscritos, a través de 144 unidades, con una capacidad instalada de 25,189 lugares, en sus tres modalidades: (i) guarderías de prestación directa por el IMSS mediante los esquemas de Madres IMSS y Ordinario, con 7 unidades para 1,407 lugares y 44 unidades para 11,370 lugares, respectivamente; y (ii) guarderías de prestación indirecta con 96 unidades y 12,412 lugares.

A pesar de los avances realizados, tan sólo el IMSS tiene un déficit de cobertura muy alto para satisfacer la demanda potencial de servicio de guardería, pero al analizar el acceso en relación a la demanda real se observa que aun hay lugares disponibles (oferta) con relación a los niños que tienen acceso a esta.

[20] Ibid

3.3 Preparación, análisis estadístico, definición y descripción de los datos

Con base en la muestra seleccionada, se realizó la estimación de la capacidad instalada en el sistema de guarderías del IMSS, tomando en cuenta la sala de atención y el tipo de modalidad, obteniéndose los datos que se presentan a continuación.

Tabla 2.1
Capacidad Instalada de las guarderías del IMSS por salas de atención, según tipo de modalidad

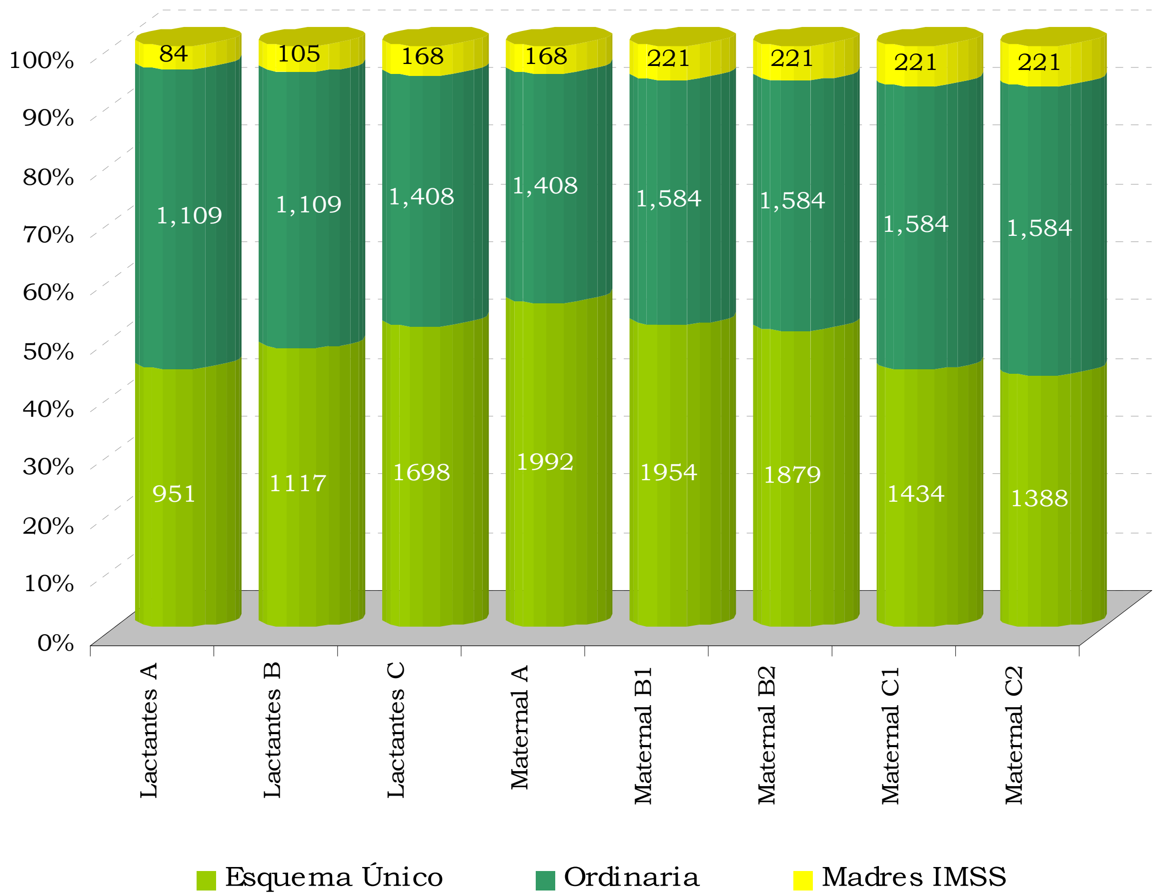
Modalidad	Ordinaria	Madres IMSS	Esquema Único	Total
Lactantes A	1,109	84	951	2,144
Lactantes B	1,109	105	1117	2,331
Lactantes C	1,408	168	1698	3,274
Maternal A	1,408	168	1992	3,568
Maternal B1	1,584	221	1954	3,759
Maternal B2	1,584	221	1879	3,683
Maternal C1	1,584	221	1434	3,238
Maternal C2	1,584	221	1388	3,193
Total	11,370	1,407	12,412	25,189

Fuente: Estimaciones por muestreo, 2008

Como se observa, el sistema de guarderías del IMSS tiene la capacidad para atender a 25,189 niños en el Distrito Federal, distribuidos en las tres modalidades, destacando el Esquema Único, con 12,412 lugares, seguido por las Ordinarias, con 11,370, y finalmente la modalidad Madres IMSS, con una capacidad de 1,407.

Con relación a la distribución de la capacidad instalada, en las modalidades y salas de atención, se muestra que es en Maternal A, B1 y B2, además del esquema único, donde se concentra la mayor capacidad de atención, con 3,568, 3,759 y 3,683 lugares, respectivamente, como se aprecia en la siguiente gráfica.

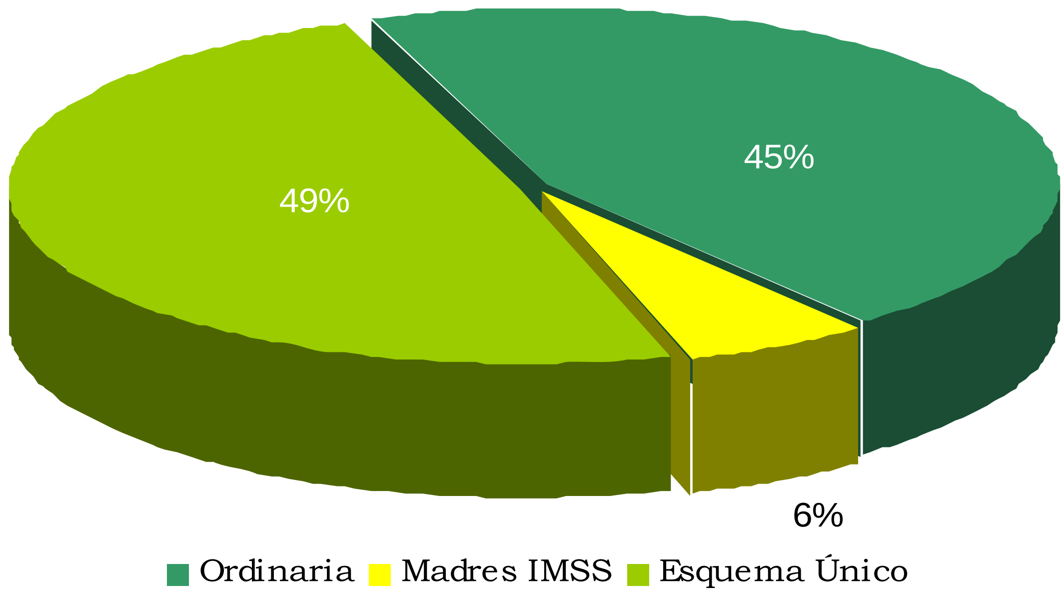
Gráfica 2.1
Distribución de la capacidad instalada de las guarderías del IMSS por salas de atención, según tipo de modalidad



Fuente: Estimaciones por muestreo, 2008

En términos porcentuales, se puede considerar que el esquema único concentra el 49% de la capacidad instalada, en tanto que las Ordinarias representan el 45%, y la modalidad Madres IMSS alcanza el 6%, lo que se refleja en la siguiente gráfica.

Gráfica 2.2
Distribución porcentual de la capacidad instalada de las guarderías del IMSS por tipo de modalidad



Fuente: Estimaciones por muestreo, 2008

Como ya se ha mencionado la capacidad instalada es mayor para ciertos rangos de edades, es decir, el servicio que dan las guarderías del IMSS es mayor a la población de niños que la utilizan en ciertas edades. En la siguiente tabla se muestra a base de estimaciones el resultado de la población total de niños inscritos en las guarderías del IMSS por edades, donde el Instituto los clasifica por salas de atención y claramente nos percatamos de que en ciertas edades la demanda potencial de los niños es mayor que en otras edades, esto se debe a factores externos que hacen que los padres que utilizan estos servicios no lleven a sus hijos.

Tabla 2.2
Promedio de la población de niños inscritos en las guarderías del IMSS por sala de atención, según tipo de modalidad

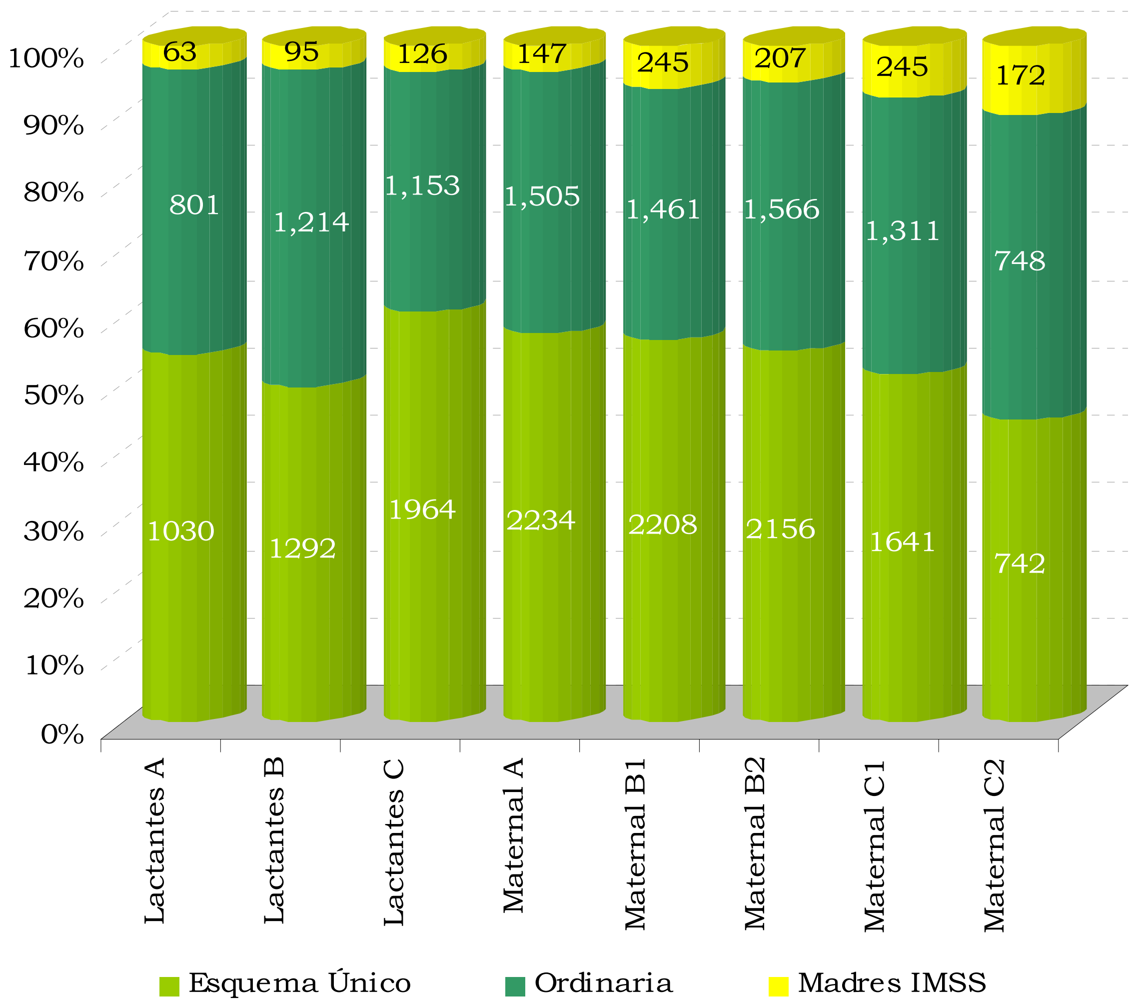
Sala de Atención	Ordinaria	Madres IMSS	Esquema Único	Total
Lactantes A	801	63	1030	1,894
Lactantes B	1,214	95	1292	2,601
Lactantes C	1,153	126	1964	3,242
Maternal A	1,505	147	2234	3,886
Maternal B1	1,461	245	2208	3,914
Maternal B2	1,566	207	2156	3,929
Maternal C1	1,311	245	1641	3,197
Maternal C2	748	172	742	1,661
Total	9,759	1,299	13,265	24,323

Fuente: Estimaciones por muestreo, 2008

Es necesario subrayar que la distribución de la población de los niños inscritos en las guarderías del IMSS es mayor en la modalidad de esquema único debido a que tienen más guarderías para dar este servicio, puesto que es más económico para el Instituto Mexicano del Seguro Social y después le siguen la modalidad ordinaria y hasta el ultimo la modalidad de Madres IMSS, ya que estas sólo cuentan con siete salas de mayor capacidad.

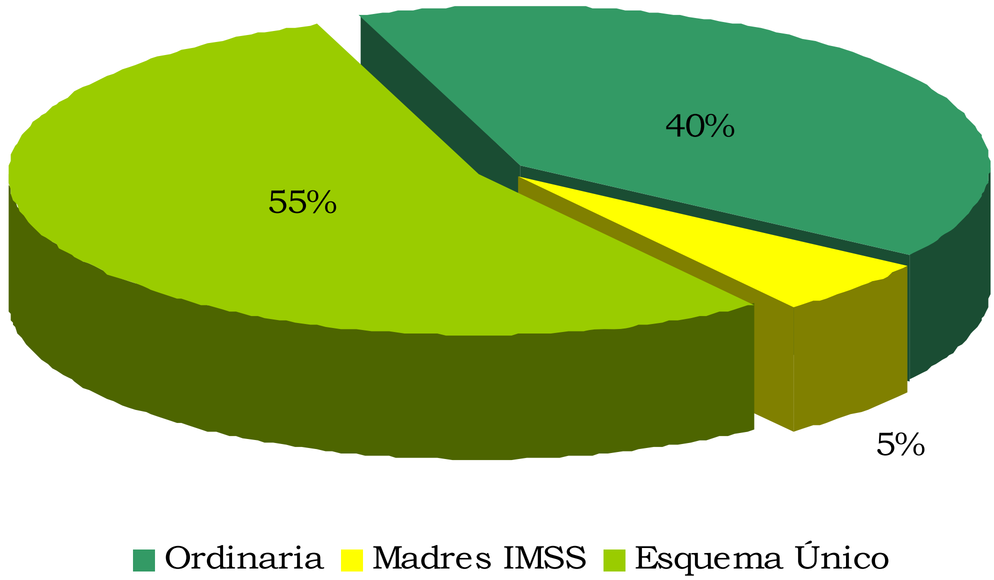
También cabe resaltar la distribución porcentual de niños inscritos por modalidad, para tener mayor claridad de lo antes dicho.

Gráfica 2.3
Distribución de la población de niños inscritos en las guarderías del IMSS por salas de atención, según tipo de modalidad



Fuente: Estimaciones por muestreo, 2008

Gráfica 2.4
Distribución porcentual de la población de niños inscritos en las guarderías del IMSS por tipo de modalidad



Fuente: Estimaciones por muestreo, 2008

Lo anterior nos conduce al tema central de la oferta de las guarderías del IMSS y la demanda potencial real de los niños. Así, al analizar el acceso que se tiene a estas guarderías, se observa que hay un diferencia, ya que el total de los lugares disponibles para los usuarios es de 25,189 comparado con el total de niños inscritos que es de 24,323, marcándose una diferencia de 866 lugares que no son ocupados, pero esto, no es del todo real, puesto que deben considerarse algunos factores externos, como por ejemplo, la edad.

En la siguiente tabla podemos apreciar que hay ciertas salas que están sobre pobladas, como aquellas que atienden a los niños que están en edades de siete meses a un año de edad (Lactantes B), los niños que tienen un año siete meses a dos años (Maternales A), los niños de edades de dos años un mes a dos años seis meses (Maternales B1), finalmente los niños que tienen edades de dos años siete meses a tres años (Maternales B2). Estas edades son las que presentan mayor demanda, aunque la capacidad instalada es mayor no alcanza cubrir estos rangos de edad.

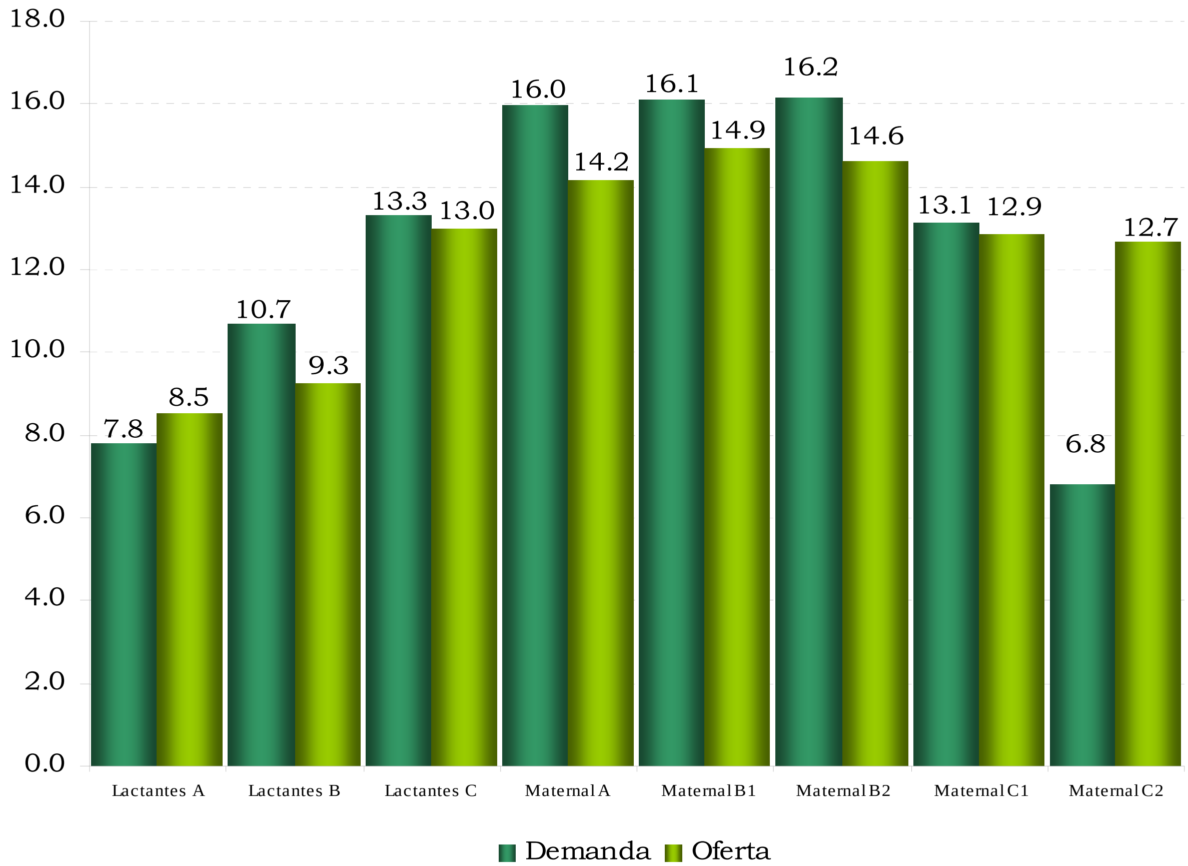
Tabla 2.3
Brecha de lugares en las guarderías del IMSS
por sala de atención

Sala de Atención	Brecha
Lactantes A	250
Lactantes B	-270
Lactantes C	31
Maternal A	-318
Maternal B1	-155
Maternal B2	-245
Maternal C1	41
Maternal C2	1,532
Total	866

Fuente: Estimaciones por muestreo, 2008

Dentro de esta afirmación, encontramos en la siguiente gráfica que la oferta sobrepasa la demanda en relación a los niños de edades de tres años a tres años seis meses y tres años seis meses a cuatro años (Maternal C2), dado que a estas edades los niños tienen la edad suficiente para inscribirse al “Kinder”, factor que determinara el deceso a la institución, por que la educación a este nivel es obligatoria. Esta es una de las dificultades que se plantean en las guarderías del IMSS, ya que no están incorporadas a la SEP.

Gráfica 2.5
Porcentaje de la capacidad instalada y niños inscritos por sala de atención (Oferta-Demanda)



Fuente: Estimaciones por muestreo, 2008

De esta manera, es preciso recalcar que en las guarderías del IMSS, hay mas niños que niñas con 50.2% y 49.8% respectivamente. La diferencia es de 0.5 por ciento, pues aquí entra el carácter de género ya que, en la actualidad siguen habiendo condicionantes culturales que nos explican que los Padres de familia prefieren que sus niñas se queden en el orden privado del hogar a que sean cuidadas en una institución público, es decir prefieren encargar su cuidado a familiares o instituciones privadas.

Tabla 2.4
Porcentaje de la población inscrita en guarderías del IMSS por sala de atención, según sexo

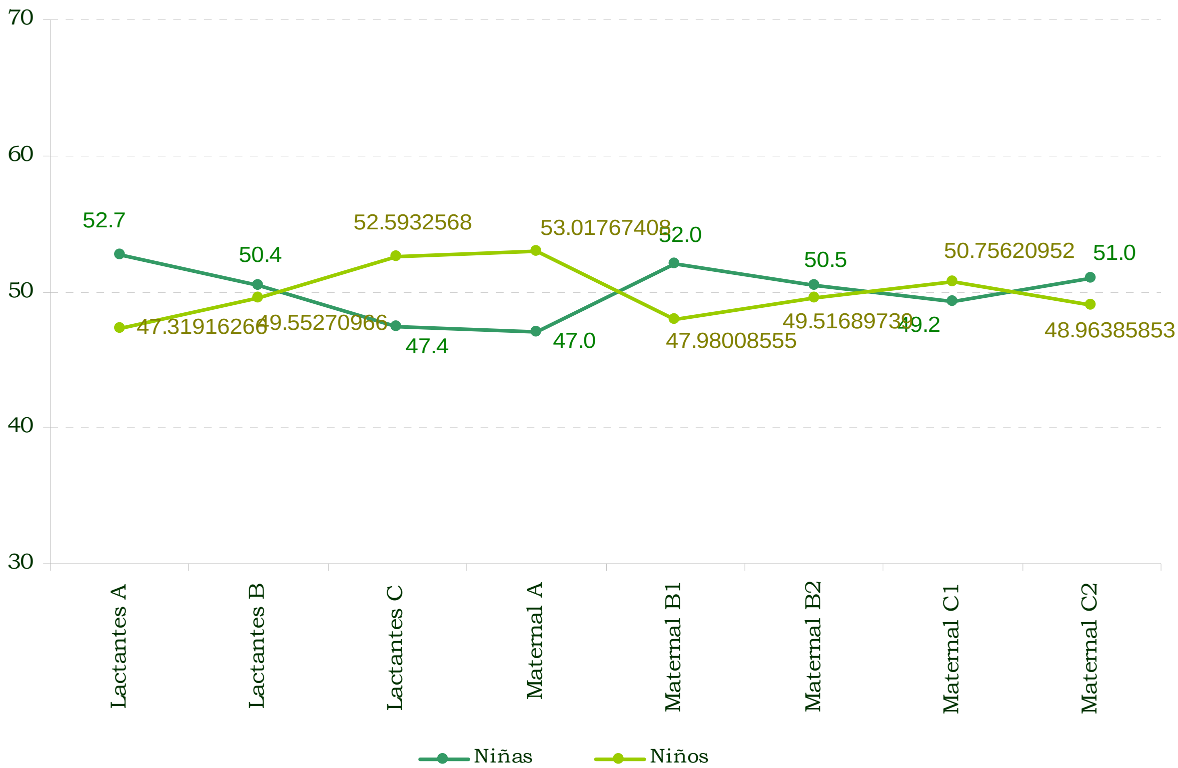
Modalidad	Niñas	Niños	Brecha
Lactantes A	52.7	47.3	5.4
Lactantes B	50.4	49.6	0.9
Lactantes C	47.4	52.6	-5.2
Maternal A	47.0	53.0	-6.0
Maternal B1	52.0	48.0	4.0
Maternal B2	50.5	49.5	1.0
Maternal C1	49.2	50.8	-1.5
Maternal C2	51.0	49.0	2.1
Total	49.8	50.2	-0.5

Fuente: Estimaciones por muestreo, 2008

De acuerdo con lo anterior dentro de los aspectos más notorios de esta división destaca el tabú en relación a la educación de la mujer puesto que es más fácil que los familiares accedan a cuidar a un bebe femenino por como fueron educados.

En la siguiente gráfica se muestra como en ciertas edades puede haber más niñas que niños pero no es suficiente para tener más población de niños en las instituciones. Pues, históricamente, el cuidado infantil ha sido tradicionalmente definido en el ámbito privado. El cuidado infantil se ha entendido generalmente como una responsabilidad privada de la familia, en realidad el cuidado infantil (y hasta la fecha) se considera responsabilidad de las madres.

Gráfica 2.6
Porcentaje de la población inscrita en las guarderías del IMSS por sala de atención, según sexo



Fuente: Estimaciones por muestreo, 2008

En cuanto al tema de los costos, la prestación de servicios de cuidado infantil, es realizada por un pequeño sector regulado. La parte del financiamiento público está dado por el patrón. De acuerdo a la ley, el patrón tiene la obligación de darle esta prestación a sus trabajadoras ya sea en su centro de trabajo o cerca de su domicilio, en base a la Ley del Seguro Social, según el artículo 211, el patrón cotiza al IMSS con la cuota del 1% por cada trabajador, en consecuencia, el costo de atención cuando es una guardería que es operada por terceros (modalidad esquema único), cuesta al instituto dos mil 556 pesos al mes por niño; cuando se trata de la modalidad de guarderías ordinarias, el IMSS paga siete mil 800 pesos al mes, debido al contrato que se tiene; y por último para la modalidad de Madres IMSS el niño por día les cuesta 18 pesos que se traduce al mes alrededor de 414 pesos.

A continuación se dan los costos aproximados que aporta el Instituto Mexicano del Seguro Social a mes por cada modalidad y cuanto aporta para cada rubro de edad de lo niños. Véase tabla 9

Tabla 9.
Costos totales mensuales para niños inscritos en las guarderías del IMSS por sala de atención, según tipo de modalidad

Modalidad	Ordinaria	Madres IMSS	Esquema Único	Total
Lactantes A	6,246,240	26,082	2,632,432	8,904,754
Lactantes B	9,472,320	39,123	3,301,694	12,813,137
Lactantes C	8,991,840	52,164	5,019,467	14,063,471
Maternal A	11,737,440	60,858	5,711,038	17,509,336
Maternal B1	11,394,240	101,430	5,644,112	17,139,782
Maternal B2	12,217,920	85,491	5,510,259	17,813,670
Maternal C1	10,227,360	101,430	4,194,043	14,522,833
Maternal C2	5,834,400	71,001	1,896,243	7,801,644
Total	76,121,760	537,579	33,909,288	110,568,627

Fuente: Estimaciones por muestreo, 2008

De esta forma es necesario evidenciar que el dinero aportado para el mantenimiento de las estancias infantiles es de orden de los cuatro mil 200 millones de pesos anuales. Es por eso, que la supervisión es bastante estricta, de hecho, las guarderías del IMSS se han ganado un gran prestigio por la atención que se da y que está relacionado con la alimentación y el resguardo de los menores.

El crecimiento de la población, empleos e inversiones propiciaron que exista cada vez un mayor número de madres trabajadoras, quienes en muchos de los casos recurren al servicio de guardería. En general, se cuentan con listas de espera y los lugares se van desocupando cuando los niños pasan a la sala siguiente de acuerdo a su edad. En la información que proporciona el IMSS, se observa, que sí hay niños en espera, es decir, hay niños que quieren entrar a las guarderías. Véase tabla 10.

Tabla 10.
Población niños en lista de espera

Año	Solicitudes Pendientes
2000	2,372
2001	2,828
2002	2,906
2003	1,666
2004	1,396
2005	1,525
2006	1,466
2007	2,446

Fuente: Instituto Mexicano del Seguro Social, 2007

Aunque el número de solicitudes pendientes parece extremadamente alto en teoría se podría suponer que las guarderías del Instituto no pueden cubrir con la demanda de atención de los niños, sin embargo, en la práctica, las guarderías sólo pueden cubrir la demanda real de los niños en ciertas edades. En las salas de atención más demandadas hay más proporción de niños en espera. Se estimaron las solicitudes pendientes y se obtuvo el promedio de los niños que están en espera. Los resultados se muestran en la siguiente gráfica y contradicen que el IMSS tenga suficientes lugares para la demanda de los niños que hay, ya que hay varios factores externos que determinan el hecho de que haya niños en espera. Un factor que lo determina, es la ubicación geográfica de las guarderías, aunque hay más guarderías concentradas en el centro no es suficiente, pues la demanda en esos lugares es demasiado alta debido a que en aquellos lugares en donde los comercios y las empresas tienen mayor presencia se espera un mayor número de derechohabientes que además de tener derecho a la prestación, requieren efectivamente hacer uso de ella. En contraparte las guarderías que están en lugares alejados de las empresas son las que tienen menor demanda de servicio para niños. Otro factor central son los aislamientos de salas (cuando a un niño le da alguna enfermedad contagiosa como varicela, rubéola, etc, el médico corresponsal de esa guardería determina por cuánto tiempo se tiene que aislar la sala, y los niños que pertenecen a esta no pueden tener contacto con los demás niños de otras salas), esto implica, que cuando un niño cumple la edad para que lo cambien de sala no pueda entrar a la que ha sido asilada entonces si hay niños que quieren entrar por primera vez a la guardería van a tener que estar en lista de espera puesto que no hay lugar. Este problema se da en varias guarderías del Distrito Federal y es un problema que deriva en una proporción alta de demanda de los niños.

Tabla 11.
Promedio de la población en lista de espera por sala de atención, según tipo de modalidad

Sala de atención	Ordinaria	Madres IMSS	Esquema único	Total
Lactantes A	138	49	137	324
Lactantes B	114	14	169	297
Lactantes C	85	4	241	330
Maternal A	114	28	177	320
Maternal B1	114	7	137	258
Maternal B2	114	4	195	313
Maternal C1	18	0	116	134
Maternal C2	0	0	111	111
Total	698	105	1,283	2,086

Fuente: Estimaciones por muestreo, 2008

Otro aspecto importante en relación con la prestación del servicio de guarderías es que los servicios asistenciales están complementados con actividades pedagógicas, donde el niño aprende a desarrollar el lenguaje, la motricidad, la sociabilización, buenos hábitos de limpieza, higiene, orden y respeto. También conocen cómo está integrada la familia, la comunidad, los medios de comunicación y de trasporte a través de los servicios pedagógicos.

De acuerdo con datos proporcionados por el IMSS, la principal actividad es la asistencial, que es su razón de ser, sin embargo, la atención del niño no sólo es darle de comer y cambiarlo, sino que durante el tiempo que permanezcan en las guarderías aprovechen para una actividad pedagógica, porque desde muy pequeños los niños deben de aprender lo que es su ropa, sus manos y cuerpo. Toda actividad asistencial da continuidad para una actividad pedagógica.

Otro aspecto notorio de este servicio es que en todas las guarderías cuentan con una sala de preparación de alimentos en la cual se muestra a los papás el desayuno y comida que se les proporcionará a sus hijos diariamente. También se guarda una muestra por si se presenta un problema de salud, para determinar la causa se analiza la porción de comida. Para la preparación de alimentos existe un instructivo de operación que se describe desde la adquisición de los víveres hasta su elaboración.

Completando el estudio, se observó que se supervisa el esquema de vacunación completa, y que su talla y peso estén de acuerdo a la edad del menor. Adicionalmente, si el menor toma leche especial, las enfermeras tienen que observar la valoración médica del niño como lo marca la normatividad y las guarderías están en todos los casos afiliadas una clínica.

Si el niño tiene alguna discapacidad, se tienen que verificar sus necesidades. En cuanto a las cuestiones administrativas las directoras tienen que estar al pendiente de tener pláticas con los papás y capacitar al personal. Aparte tienen supervisiones del IMSS.

Además, en lo que compete al personal, se hace un proceso selectivo mediante un examen psicológico, deben tener un perfil académico para el puesto que van a desempeñar y posteriormente tienen que realizar prácticas en la guardería por lo menos dos semanas y al pasar ese proceso se les hace una examen médico y procede la contratación, antes no se puede tener ese recurso humano en las guarderías. Las guarderías de la subdelegación son de particulares que otorgan la atención exclusiva al Seguro Social. El prestador está sujeto a darle servicio mediante recurso técnico que el IMSS establece, y esos recursos técnicos son administración, pedagogía, fomento a la salud y nutrición.

CONCLUSIONES

El IMSS enfrenta el gran reto de no sólo crecer al ritmo que las transformaciones en materia de igualdad de género exigen, sino con mayor celeridad para satisfacer la demanda potencial

Un resultado lógico del cuidado infantil se ve altamente condicionado por el acceso, la disponibilidad económica y la calidad que ofrecen. Pese a que la demanda aún no cubre la capacidad máxima, el servicio crece, pues al analizar las cifras, se advierte que el sistema de guarderías públicas no cumple con el propósito de atender a toda la población entre los cero y cuatro años de edad, pues la baja proporción de guarderías públicas se pone en evidencia si se toman en cuenta dos factores adicionales: el primero , consiste en que se ha incrementado el número de mexicanas dentro de la población Económicamente Activa y, el segundo, relativo a la existencia de hogares catalogados como monoparentales, los cuales requieren de un mayor apoyo.

En otros caso, y es materia fundamental en este estudio, las guarderías no están cerca del lugar de trabajo o en caso de que las hubiera, se encuentran saturadas. Por ejemplo, en el Distrito Federal las guarderías que están en el centro de la ciudad cuentan con mayor población de infantes en ciertos rangos de edad, por lo tanto, hay mucho más población infantil en espera. En cuánto a la ubicación, una forma ideal para abatir este problema, sería abrir más guarderías donde se concentra más la población trabajadora, exhortando al Estado que preste servicios de guardería, sin perjuicio de que el sector público los brinde de forma directa, de buena calidad, flexibles y asequibles.

Históricamente, el ramo de guarderías ha presentado un rezago en cuanto a la cobertura potencial de servicios que podría demandar la población con derecho a esta prestación. Como ya se ha demostrado la demanda potencial de los niños es mayor, pues el IMSS sólo atiende la quinta parte de la población de infantes, sin embargo en este estudio se demuestra que la demanda potencial real de los infantes sí se satisface puesto que hay lugares disponibles. Pero aún así, se prevé, que la demanda real va a superar a la oferta, es por eso que se tienen que tomar medidas.

Lo que se propone en base a este estudio es que se tenga la posibilidad de abrir mayores espacios de atención en las siguientes edades de siete meses a un año (Lactantes B); un año siete meses a dos años (Maternales A); dos años un mes a dos años seis meses (Maternales B1); finalmente dos años siete meses a tres años (Maternales B2), puesto que hay mucha demanda y que en la sala de atención de Maternales C2 se transfieran las salas que tienen más demanda puesto a que en estas edades la población tiene a ser menor, ya que a esta edad, se tienen que inscribir en el “Kinder”.

Con base en lo antes expuesto, es preciso señalar que las guarderías del IMSS prestan un servicio de calidad y efectivamente están comprometidas con la salud del niño, así como su desarrollo familiar, social, intelectual y conductual. Es difícil dar este servicio debido a la irresponsabilidad que algunos padres tienen, aunque existen normas y reglas que regulan la responsabilidad de ambas partes se observó que no se lleva acabo por parte de los padres, esto repercute gravemente en el costo de las guarderías, dado que si los niños faltan, no se les paga ese día, como se analizó. En promedio, los que niños faltan son 5,048 al día, esto quiere decir que, en promedio, las guarderías pierden 1,244,451 dinero al día, y puesto que las guarderías llevan un rol de administración semanal para la alimentación de los infantes esto provoca pérdida monetaria, pues si se hace la despensa cada semana, no se sabe cuantos niños van a faltar y aún así se tiene que comprar la despensa necesaria. Cabe señalar que en algunas guarderías los padres son comprometidos, pero este sector es minoría y es por eso que se debe concientizar a los padres de familia que cumplan con las normas impuestas par el buen funcionamiento de estas guarderías y así no se pierdan recursos.

Un punto importante como parte de este proceso, es que los recursos destinados a ese servicio no llegaban a su destino manifiesto porque, como tantas otras cosas, se desviaban para atender el seguro de enfermedades generales y maternidad, que son parte de la prestación de servicios médicos, protagonista importante de la vida del IMSS. Pero al mismo tiempo, con el crecimiento de la población, los servicios de guardería tenían que atenderse sin la posibilidad económica de hacerlo directamente el IMSS, lo que es cierto es que hace ya un número importante de años, el sistema de guarderías vive bajo el impulso económico de los particulares y la vigilancia estrecha del IMSS, dando a notar que la cuota de del seguro social de guarderías es particularmente baja .

Otro importante argumento en relación con la prestación de servicios de cuidado infantil se realiza por razones de carácter redistributivo para el país convirtiéndose en una contribución a la ciudadanía y es particularmente importante para la reducción de la pobreza de las mujeres. El cuidado infantil de calidad es un factor que contribuye a la cohesión e inclusión social, el aprendizaje de a lo largo de la vida y al tipo de experiencias de la temprana infancia que ayudan a la construcción de la potencialidad en convertirse en un servicio de ayuda a la construcción de comunidades cuidadoras, el aprendizaje temprano y el cuidado mejora los resultados del niño, por ejemplo, la preparación escolar, aspecto relativamente fácil de medir, guarda correlación con un buen cuidado infantil, y parece estar fuertemente asociado a un mejor desempeño escolar y a tasa más bajas de deserción.

BIBLIOGRAFIA

- Monografias: Conceptos Básicos de Muestreo. Ignacio Méndez Ramírez, Guillermina Eslava y Patricia Romero Mares, Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, UNAM.
- EL PROCESO DE LA INVESTIGACIÓN CIENTÍFICA, Mario Tamayo y Tamayo, 1998.
- TECNICAS DE MUESTREO, William G. Cochran, 1984.
- CURSO INTENSIVO DE MUESTREO EN POBLACIONES FINITAS, J. L. Sánchez Crespo, Instituto Nacional de Estadística, Madrid 1980.
- Notas del Profesor Francisco Sánchez Villareal, muestro 2005.
- IMSS. Ley del Seguro Social
- IMSS. Informe al Ejecutivo Federal y al Congreso de la Unión sobre la situación financiera y los riesgos del Instituto Mexicano del Seguro Social, 2005-2006
- IMSS. Informe de la Evaluación de Riesgos Contenidos en Programa de Administración de Riesgos Institucionales (PARI) 2005-2006.
- CONAPO. Proyecciones de la población de México 2000 - 2050. México. 2002.
- INEGI. Indicadores sociodemográficos de México (1930-2000).

ANEXOS

Anexo 1. Muestreo Aleatorio de Conglomerados o Muestreo Polietápico

Un conglomerado es un conjunto de elementos de la población, de donde una muestra por conglomerados es una muestra aleatoria en la cual cada unidad muestral es una colección o conglomerado de elementos.

Este diseño de muestra es útil, cuando la población contiene muchas unidades más o menos dispersas, puede haber dos razones fundamentales que impidan la toma de muestras directamene de la población:

- No se dispone con marco de elementos para las unidades de la población, o son muy caros o imposible de construirlo.
- El costo del muestreo se incrementa mucho por la dispersión de la población de las unidades, con lo que se desea obtener la muestra menos dispersa, es menos costoso que el muestreo aleatorio simple.

Cuando sucede algún evento antes mencionado, o lo que es común, que ocurra de manera simltánea, se recomienda el uso del muestreo por conglomerados, aunque puede haber posibilidad de producir varianzas mayores que un muestreo directo de una etapa.

En algunas situaciones el tamaño de los conglomerados (número de elementoa que lo componen), está dado y en otros casos se tiene que definir el tamaño dependiendo de las circuntancias que se quiere medir.

Con Muestreo de Conglomerados

En el muestreo por conglomerados queremos que los conglomerados contengan unidades muy heterogéneas dentro de estos y homogéneas entre ellos.

Definición y notación

A nivel poblacional:

N	Números de conglomerados en la población.
n	Número de conglomerados en la muestra.
M_i	Número de elementos en el conglomerado $i, i=1,\dots,N$.
$M = \sum_{i=1}^{N} M_i$	Total de elementos en la población.
$Y_{i,j}$	Valor de la medición de elementos j del conglomerados i (a veces no la tenemos).
$Y_i = \sum_{i=1}^{M_i} Y_{ij}$	Total de conglomerado i (a veces es lo que tenemos).
$\bar{Y}_i = \frac{1}{M_i} \sum_{i=1}^{M_i} Y_{ij}$	Promedio del conglomerado i.

$$Y = \sum_{i=1}^{N} Y_i = \sum_{i=1}^{N}\sum_{j=1}^{M_i} Y_{ij}$$ Total poblacional.

$$\overline{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i$$ Promedio de totales de conglomerados.

$$\overline{Y}_e = \frac{Y}{M} = \frac{\sum_{i=1}^{N} Y_i}{\sum_{i=1}^{M} M_i}$$ Promedio por elemento.

$$S_b^2 = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \overline{Y})^2$$ Varianza entre totales de conglomerados.

Estimadores y varianza

Estos dos estimadores, el del total poblacional y de la media poblacional por elemento, son insesgados , pero frecuentemente tienen varianzas grandes, ya que si el número de elementos en los conglomerados M_i, es muy diferente, genera variabilidad entre los totales de los conglomerados.

Si tenemos una muestra aleatoria simple de n conglomerados.
El estimador del promedio por conglomerado es:

$$\hat{\overline{Y}} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

donde $y_i = \sum_{i=1}^{M_i} y_{ij}$ es el total observado del conglomerado *i.*
El estimador del total poblacional Y es:

$$\hat{Y} = N\hat{\overline{Y}} = \frac{N}{n}\sum_{i=1}^{n} y_i = \frac{N}{n}\sum_{i=1}^{n}\sum_{j=1}^{M_i} y_{ij}$$

Con varianza y estimador de varianza:

$$V(\hat{Y}) = N^2\left(1 - \frac{n}{N}\right)\frac{S_b^2}{n}$$

$$\Rightarrow \qquad \hat{V}(\hat{Y}) = N^2\left(1 - \frac{n}{N}\right)\frac{\hat{S}_b^2}{n}$$

donde,

$$\hat{S}_b^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \hat{\overline{Y}})^2$$

Estimador de la media poblacional por elemento, si se conoce *M*, el total de elementos de la población, entonces, el estimador de la media poblacional por elemento es:

$$\bar{Y}_e = \frac{Y}{M} = \frac{\sum_{i=1}^{N} Y_i}{\sum_{i=1}^{M} M_i}$$

Con varianza y estimador de varianza:

$$V(\hat{\bar{Y}}_e) = \frac{1}{M^2} V(\hat{Y})$$

$$V(\hat{\bar{Y}}_e) = \frac{1}{M^2} \hat{V}(\hat{Y})$$

Si el tamaño del conglomerado M_i está fuertemente relacionado con el total del conglomerado, lo que generalmente sucede, entonces se prefieren los estimadores de razón.

Estimador de la media poblacional por elemento de razón

$$\hat{\bar{Y}}_e = \frac{\hat{Y}}{\hat{M}} = \frac{\frac{N}{n}\sum_{i=1}^{n} y_i}{\frac{N}{n}\sum_{i=1}^{n} M_i} = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} M_i}$$

Con varianza:

$$V(\hat{\bar{Y}}_e) = \left(1 - \frac{n}{N}\right)\frac{1}{n}\frac{1}{\bar{M}^2}\sum_{i=1}^{N}\frac{(Y_i - \bar{Y}_e M_i)^2}{N-1}$$

Donde

$\bar{M} = \frac{M}{N}$ es el tamaño promedio de los conglomerados

Estimador de varianza:

$$V(\hat{\bar{Y}}_e) = \left(1 - \frac{n}{N}\right)\frac{1}{n}\frac{1}{\hat{\bar{M}}^2}\sum_{i=1}^{N}\frac{(Y_i - \bar{Y}_e M_i)^2}{N-1}$$

donde,

$$\hat{\bar{M}} = \frac{\hat{M}}{N} = \frac{\frac{N}{n}\sum_{i=1}^{n} M_i}{N} = \sum_{i=1}^{n}\frac{M_i}{n}$$

Estimador del total poblacional con razón.

$\hat{Y} = M\hat{\bar{Y}}_e$ con M conocida

Con estimador y varianza:

$$V(\hat{Y}) = M^2 V(\hat{\bar{Y}}_e)$$

$$V(\hat{Y}) = M^2 \hat{V}(\hat{\bar{Y}}_e)$$

Estimador de una proporción poblacional con razón

Sea

$$Y_{ij} = \begin{cases} 1 & U_{ij} \textit{ tiene la característica} \\ 0 & U_{ij} \textit{ no tiene la característica} \end{cases}$$

El estimador de la proporción de unidades con la característica es:

$$\hat{P} = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} M_i}$$

Con varianza estimada:

$$\hat{V}(\hat{P}) = \left(1 - \frac{n}{N}\right)\frac{1}{n}\frac{1}{\hat{\bar{M}}}\sum_{i=1}^{n}\frac{(y_i - \hat{P}M_i)^2}{n-1}$$

Determinación de la Muestra

Se fijan la precisión δ y la confianza (1-α)

$$\delta = Z_{(1-\alpha/2)}\sqrt{V(\hat{\bar{Y}}_e)}$$

$$\delta^2 = Z_{(1-\alpha/2)}\left(\frac{1}{n} - \frac{1}{N}\right)\frac{1}{\bar{M}^2}S_b^2$$

Despejando *n* para obtener el tamaño de muestra tenemos que:

$$n = \frac{NZ_{(1-\alpha/2)}S_b^2}{N\delta^2\bar{M}^2 + Z_{(1-\alpha/2)}S_b^2} = \frac{NS_b^2}{\frac{N\delta^2\bar{M}^2}{Z_{(1-\alpha/2)}} + S_b^2}$$

Anexo 2. Muestreo con censo de Conglomerados o Bietápico

No se censan los conglomerados en muestra, si no que se toma una muestra de sus elementos. El principio básico para construir estimadores y sus varianzas en muestreo bietápico, es considerar el muestreo de unidades secundarias dentro de cada unidad primaria y obtener estimadores de totales de *Y,* y sus varianzas en las UPM. Después usando los estimadores para cada unidad primaria muestreada, es necesario considerar un muestreo de unidades primarias, con el fin de estimar totales en toda la población
Se selecciona al azar (m.a.s.)

Definición y Notación

A nivel Poblacional:

N Número de unidades primarias de muestreo (UPM). Se cuenta con un Marco de muestreo para las N UPM.

M_i Número de unidades secundarias (USM) en la unidad primaria *i*-ésima.

$M = \sum_{i=1}^{N} M_i$ Número total de (USM). Que normalmente no se conoce

Y_{ij} Valor de la medición en la unidad secundaria *j*-ésima dentro de la unidad *i*-ésima.

$Y_i = \sum_{j=1}^{M_i} Y_{ij}$ Total de los valores de Y_{ij} de la unidad primaria *i*-ésima.

$\bar{Y}_i = \frac{1}{M_i} Y_i = \frac{1}{M_i} \sum_{j=1}^{M_i} Y_{ij}$ Promedio de la unidad primaria *i*-ésima.

$Y_i = M_i \bar{Y}_i$ Total de la unidad primaria *i*-ésima.

$Y = \sum_{i=1}^{N} Y_i = N\bar{Y} = \sum_{i=1}^{N} \sum_{j=1}^{M_i} Y_{ij}$ Total de los valores de la característica de Y en la población.

$\bar{Y} = \frac{1}{N} \sum_{i=1}^{N} Y_i$ Promedio de los totales de unidades primarias, este promedio de totales es completamente diferente al promedio por elemento.

$\bar{Y}_e = \frac{Y}{\sum_{i=1}^{N} M_i} = \frac{Y}{M}$ Promedio por elemento.

$S_{wi}^2 = \frac{\sum_{j=1}^{M_i} (Y_{ij} - \bar{Y}_i)^2}{M_i - 1}$ Varianza entre unidades secundarias dentro de la unidad primaria *i*-ésima.

$S_b^2 = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \bar{Y})^2$ Varianza entre totales de conglomerados.

A nivel muestral:

n	Número de UPM en muestra.
m_i	Número de USM muestreadas en la UPM$_i$.
y_{ij}	Medición de la USM$_j$ en muestra de la UPM$_i$.
$\hat{\bar{Y}}_i = \bar{y}_i = \frac{1}{m_i}\sum_{j=1}^{m_i} y_{ij}$	Promedio muestral de las USM de la UPM$_i$.
$\hat{Y}_i = M_i\hat{\bar{Y}}_i$	Promedio estimado de la UPM$_i$.
$\hat{S}_{wi}^2 = \frac{1}{m_i - 1}\sum_{j=1}^{m_i}(y_{ij} - \bar{y}_i)^2$	Varianza estimada entre USM dentro de las UPM$_i$.
$\bar{\hat{Y}} = \frac{1}{n}\sum_{i=1}^{n}\hat{Y}_i = \frac{1}{n}\sum_{i=1}^{n}M_i\hat{\bar{Y}}_i$	Promedio de totales estimados de la UPM.

Estimadores y varianza

Para construir estimadores, se considera el muestreo aleatorio simple, para UPM y también muestreo aleatorio simple para USM., aunque puede haber otras formas de tomar la muestra en ambas etapas.

Total estimado para la población es:

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

$$\Rightarrow \qquad = \frac{N}{n}M_i\bar{y}_i = \frac{N}{n}\sum_{i=1}^{n}M_i\frac{1}{m_i}\sum_{j=1}^{m_i}y_{ij}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\underbrace{\frac{N}{n}\frac{M_i}{m_i}}_{\substack{\text{Factores de} \\ \text{expansión } f_i}} y_{ij} = \sum_{i=1}^{n}\sum_{j=1}^{m_i}f_{ij}y_{ij} = N\frac{\sum_{i=1}^{n}\hat{Y}_i}{n}$$

Si m_i es proporcional a M_i entonces $\frac{M_i}{m_i} = k$ el diseño es autoponderado, es decir los factores de expansión son iguales, $f_{ij} = f = \frac{N}{n}k$, $\forall i, \forall j$.

El estimador de la varianza total es:

$$V(\hat{Y}) = \underbrace{N^2\left(\frac{1}{n} - \frac{1}{N}\right)S_b^2}_{1} + \frac{N}{n}\sum_{i=1}^{N}M_i^2\underbrace{\left(\frac{1}{m_i} - \frac{1}{M_i}\right)}_{2}S_{wi}^2$$

(1)Es el 90%-95% del valor de $V(\hat{Y})$.

(2)Es cero si $m_i = M_i$, es decir, si se censan las UPM. Es el caso del muestreo por conglomerados.

Es común que los valores de Y_{ij} sean semejantes dentro de las UPM. Esto hace que las S_{wi}^2 tiendan a ser pequeñas y los valores de las Y_{ij} son más diferentes entre UPM; además, los totales Y_i difieren mucho si los tamaños M_i de USM, dentro de las UPM son diferentes mientras que S_b^2 se tiene varianza entre totales, no entre valores individuales. Todo esto hace que la primera parte de $V(\hat{Y})$ constituya con frecuencia un 90%-95% o incluso más valor de total de la varianza del estimador total.

Cuando los valores de las Y_{ij} tienden a ser parecidos dentro de cada una de las UPM, entonces se genera una correlación, llamada correlación intraconglomerado, esta correlación hace que la información tenga cierta redundancia, lo que se refleja en la varianza de los estimadores mayor que ka que se obtendría en un muestreo unietápico de las unidades.

La varianza del estimador del total se estima con:

$$V(\hat{Y}) = N^2\left(\frac{1}{n} - \frac{1}{N}\right)\hat{S}_b^2 + \frac{N}{n}\sum_{i=1}^{N} M_i^2\left(\frac{1}{m_i} - \frac{1}{M_i}\right)\hat{S}_{wi}^2$$

Donde:

$$S_b^2 = \frac{1}{n-1}\sum_{i=1}^{n}(\hat{Y}_i - \bar{\hat{Y}})^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left[M_i\hat{\bar{Y}}_i - \frac{1}{n}\sum_{i=1}^{n} M_i\bar{y}_i\right]$$

El intervalo aproximado del $(1-\alpha)100\%$ de confianza para Y será:

$$P\left[\hat{Y} - Z_{(1-\alpha/2)}\sqrt{\hat{V}(\hat{Y})} < Y < \hat{Y} + Z_{(1-\alpha/2)}\sqrt{\hat{V}(\hat{Y})}\right] = 1-\alpha$$

Estimador de la media por elemento es:

$$\hat{\bar{Y}}_e = \frac{\hat{Y}}{\hat{M}} = \frac{\frac{N}{n}\sum_{i=1}^{n}\hat{Y}_i}{\frac{N}{n}\sum_{i=1}^{n}M_i} = \frac{\sum_{i=1}^{n}\hat{Y}_i}{\sum_{i=1}^{n}M_i} = \frac{\sum_{i=1}^{n}M_i\bar{y}_i}{\sum_{i=1}^{n}M_i}$$

Varianza del estimador de la media por elemento será:

$$V(\hat{\bar{Y}}_e) = \left(1 - \frac{n}{N}\right)\frac{1}{n\bar{M}^2}\sum_{i=1}^{N}\frac{M_i^2(\bar{Y}_i - \bar{Y}_e)^2}{N-1} + \frac{1}{nN\bar{M}^2}\sum_{i=1}^{N}M_1^2\left(1 - \frac{m_i}{M_i}\right)\frac{S_{wi}^2}{m_i}$$

Con estimador:

$$\hat{V}(\hat{\bar{Y}}_e) = \left(1 - \frac{n}{N}\right)\frac{1}{n\hat{\bar{M}}^2}\sum_{i=1}^{N}\frac{M_i^2(\bar{Y}_i - \bar{Y}_e)^2}{N-1} + \frac{1}{nN\hat{\bar{M}}^2}\sum_{i=1}^{N}M_1^2\left(1 - \frac{m_i}{M_i}\right)\frac{\hat{S}_{wi}^2}{m_i}$$

Estimador de la media por elemento, si se conoce *M,* el total de *USM* en la población, otra forma de estimar la media por elemento es:

$$\hat{\bar{Y}}_e = \frac{\hat{Y}}{M} = \frac{N}{nM}\sum_{i=1}^{n}M_i\bar{y}_i \ ,$$

Con varianza y estimador de varianza:

$$V(\hat{\bar{Y}}_e) = \frac{1}{M^2} V(\hat{Y})$$

$$\Rightarrow \qquad \hat{V}(\hat{\bar{Y}}_e) = \frac{1}{M^2} \hat{V}(\hat{Y})$$

Estimador de una proporción es:

$$\hat{P} = \frac{\sum_{i=1}^{n} M_i \hat{p}_i}{\sum_{i=1}^{n} M_i}$$

donde, $\hat{p}_i$ es la proporción en la *UPM*$_i$, entonces el estimador de la varianza del estimador de la proporción es:

$$\hat{V}(\hat{P}) = \left(1 - \frac{n}{N}\right) \frac{1}{n\hat{\bar{M}}^2} \sum_{i=1}^{n} \frac{M_i^2 (\hat{p}_i - \hat{P})^2}{n-1} + \frac{1}{nN\hat{\bar{M}}^2} \sum_{i=1}^{N} M_1^2 \left(1 - \frac{m_i}{M_i}\right)\left(\frac{\hat{p}_i (1 - \hat{p}_i)}{m_i}\right)$$

Determinación del tamaño de Muestra

Si se conoce el primer término de $V(\hat{Y})$ representa la mayor parte (95% o más) de su valor, es decir, se desprecia la variación entres *USM* dentro de *UPM,* se tiene:

$$\delta = Z_{(1-\alpha/2)} \sqrt{V(\hat{Y})} = Z_{(1-\alpha/2)} \sqrt{N^2 \left(\frac{1}{n} - \frac{1}{N}\right) S_b^2}$$

Despejando *n:*

$$\Rightarrow \qquad n = \frac{1}{\dfrac{\dfrac{\delta^2}{Z_{(1-\alpha/2)} S_b^2}}{} + \dfrac{1}{N}} \doteq \frac{N Z_{(1-\alpha/2)} S_b^2}{N\delta^2}$$

Anexo 3. Cuestionario

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACUTLAD DE CIENCIAS

DEPARTAMENTO DE MATEMÁTICAS

Cuestionario de Guarderías del IMSS

Datos Generales

Modalidad

Responsable

Nombre(s) Apellido(s)

Cargo / Puesto

Domicilio

calle No.

Colonia Municipio / Delegación

Teléfono **Correo - e**

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
FACUTLAD DE CIENCIAS
DEPARTAMENTO DE MATEMÁTICAS

Cuestionario de Guarderías del IMSS

1.- Lactantes A

Inscritos		**En trámite**	
Total		**Total**	
H		**H**	
M		**M**	
Lugar de residencia		**Capacidad instalada**	
DF		**Total**	
Edo. Méx.		**costo del servicio**	
Otros			

2.- Lactantes B

Inscritos		**En trámite**	
Total		**Total**	
H		**H**	
M		**M**	
Lugar de residencia		**Capacidad instalada**	
DF		**Total**	
Edo. Méx.		**costo del servicio**	
Otros			

3.- Lactantes C

Inscritos		**En trámite**	
Total		**Total**	
H		**H**	
M		**M**	
Lugar de residencia		**Capacidad instalada**	
DF		**Total**	
Edo. Méx.		**costo del servicio**	
Otros			

4.- Maternal A

Inscritos
Total
H
M

Lugar de residencia
DF
Edo. Méx.
Otros

En trámite
Total
H
M

Capacidad instalada
Total
costo del servicio

5.- Maternal B1

Inscritos
Total
H
M

Lugar de residencia
DF
Edo. Méx.
Otros

En trámite
Total
H
M

Capacidad instalada
Total
costo del servicio

6.- Maternal B2

Inscritos
Total
H
M

Lugar de residencia
DF
Edo. Méx.
Otros

En trámite
Total
H
M

Capacidad instalada
Total
costo del servicio

7.- Maternal C1

Inscritos
Total
H
M

Lugar de residencia
DF
Edo. Méx.
Otros

En trámite
Total
H
M

Capacidad instalada
Total
costo del servicio

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
FACUTLAD DE CIENCIAS
DEPARTAMENTO DE MATEMÁTICAS

Cuestionario de Guarderías del IMSS

8.- Maternal C2

Inscritos

Total
H
M

Lugar de residencia

DF
Edo. Méx.
Otros

En trámite

Total
H
M

Capacidad instalada

Total

costo del servicio

Titulares de derechohabiencia

Total
H
M

Edades

15 a 19
20 a 24
25 a 29
30 a 34
35 a 39
40 y más

Trabajadores
Médicos

Total
Hombres
Mujeres

Enfermeras

Total
Hombres
Mujeres

Otros

Total