



UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO

FACULTAD DE CIENCIAS

EL CAMBIO EN LOS DETERMINANTES DE LA
MORTALIDAD INFANTIL EN MÉXICO
1987-1997

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

ACTUARIA

P R E S E N T A :

María Felipa Hernández López

TUTOR

Mat. Javier González Rosas

2007





Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

Datos del Jurado

<p>1. Datos del alumno Hernández López María Felipa 59 75 01 14 Universidad Nacional Autónoma de México Facultad de Ciencias Actuaría 092518577</p>
<p>2. Datos del tutor Mat Javier González Rosas</p>
<p>3. Datos del sinodal 1 M en D María Teresa Velázquez Uribe</p>
<p>4. Datos del sinodal 2 Dra María Edith Pacheco Gómez Muñoz</p>
<p>5. Datos del sinodal 3 M en P Laura Elena Gloria Hernández</p>
<p>6. Datos del sinodal 4 M en D Miguel Ángel Martínez Herrera</p>
<p>7. Datos del trabajo escrito El Cambio en los Determinantes de la Mortalidad Infantil en México 1987-1997 78 p. 2007</p>

Agradecimientos

Con todo mi amor a mi abuelita Felipa Martínez Rodríguez, porque gracias a su apoyo logro culminar una etapa de mi vida

A mis tías Guadalupe, Linda y Ángela y mi tío Elías por estar siempre conmigo

A mis primas Irma y Elsa por ser mis hermanas

A mis hijas Dayra y Brenda que son la luz y amor que me da la fuerza para seguir adelante

A Yolanda por brindarme su amistad y ser mi amiga en todo momento

A Xavier por toda la paciencia y ayuda incondicional para la realización de este trabajo, sinceramente no tengo palabras para agradecerte,
GRACIAS

Índice

Introducción	I
Capítulo I. Marco Teórico conceptual de la mortalidad infantil	
1.1. Definición de Muerte.....	1
1.2. Determinantes de la mortalidad.....	3
1.3. Determinantes próximos.....	5
1.4. Determinantes subyacentes.....	7
1.5. El patrón de formación familiar.....	8
1.6. La mortalidad infantil y el patrón de formación familiar.....	11
Capítulo II. El modelo de regresión logística	
2.1. Tipo de variables.....	16
2.2. Regresión logística simple.....	18
2.3. Regresión logística múltiple.....	27
2.4. Estimación de parámetros.....	27
2.5. Validación del modelo logístico.....	33
2.6. Bondad de ajuste del modelo.....	36
Capítulo III. Determinantes de la mortalidad en México	
3.1. Tamaño de muestra.....	41
3.2. Análisis comparativo de las variables del patrón familiar.....	43
3.3. Análisis del patrón de formación familiar desde el enfoque ordinal.....	47
3.4. Variables biosocioeconómicas.....	49
3.5. Análisis multivariado de la mortalidad infantil.....	54
3.6. Validación del modelo.....	57
Resumen y conclusiones.....	61
Anexo.....	67
Bibliografía.....	69
Guía de cuadros y gráficas.....	73

I

Introducción

En México, la mortalidad infantil ha sido uno de los principales puntos de reflexión y atención de los investigadores y políticos y fue hasta los años treinta cuando se dio una transición que marcó el descenso de los niveles, como producto de la inversión e importación de infraestructura en nuestro país.

El rápido descenso de la mortalidad, se ha debido principalmente a los avances tecnológicos en la prevención y control de enfermedades y a la expansión de los servicios médicos y de salud pública, que son independientes de las condiciones socioeconómicas de cada región y que en la actualidad inciden de manera importante en la mortalidad de los niños y niñas durante el primer año de su vida.

En las últimas décadas, la mortalidad infantil ha descendido paulatinamente hasta alcanzar cifras aproximadas a diecinueve defunciones por cada mil nacidos vivos¹ en el año 2000, lo que refleja que aún es grave el problema de la mortalidad infantil en nuestro país. A este respecto, el presente trabajo tiene como punto central responder a las siguientes preguntas:

- ¿Se asocian la mortalidad infantil y el patrón de formación familiar?
- ¿Cuáles han sido los cambios que se han dado en el patrón de formación familiar entre un periodo y otro?

¹ CONAPO, (2002). Proyecciones de población 2000-2030, México.

- ¿Cuáles son las principales variables sociales, económicas y demográficas que se relacionan con la mortalidad infantil?

El estudio se deriva, principalmente, de dos investigaciones sobre mortalidad infantil, la primera; *La Influencia de los Patrones de Formación Familiar en la Mortalidad Infantil* de José Luis Bobadilla, con datos de la Encuesta Nacional de Fecundidad 1987 (ENFES), misma que tiene sustento en el estudio de Hobcraft, McDonald y Rutstein (1985); la segunda, se refiere a *Timing of Births and Child Survival* de Naciones Unidas (1994), en el cual se hace una comparación entre 25 países, sobre la influencia de los patrones de formación familiar y otras variables socioeconómicas de la mortalidad en los niños.²

Ambos estudios se realizaron para 1987, por lo que, la importancia de este trabajo radica en analizar, los cambios que se han presentado durante la década de 1987 a 1997, con base en la información sobre la historia de embarazos, captada en la Encuesta Nacional de la Dinámica Demográfica de 1997 (ENADID).

El trabajo está desarrollado en tres capítulos, en el primer capítulo de este trabajo, se hace referencia a la definición de muerte desde el punto de vista filosófico y médico, a los determinantes de la mortalidad, así como a los principales resultados sobre la mortalidad infantil a partir de la conformación del patrón de formación familiar, de los estudios realizados por Hobcraft, McDonald y Rutstein y José Luis Bobadilla. En el capítulo segundo describen los principales planteamientos teóricos del modelo de regresión logística, el cual se utiliza para realizar el análisis estadístico de la influencia simultánea de los diferentes factores sociales, económicos y demográficos, que influyen en la mortalidad infantil. Finalmente, en el capítulo tercero se presentan los resultados descriptivos de la mortalidad infantil con respecto al patrón de formación familiar y con cada una de las variables socioeconómicas, así como los resultados de la aplicación del modelo de regresión logística en los dos periodos de análisis.

² Se han realizado diversos estudios referentes al análisis de la mortalidad infantil, pero en particular estos dos estudios se basan en la misma línea de análisis.

CAPÍTULO I

Marco teórico conceptual de la mortalidad infantil en México

1.1 Definición de muerte¹

El vocablo TANATOS se deriva del griego THANATOS, que era el nombre que se le daba a la diosa de la muerte en la mitología griega. LOGOS se deriva también del griego, que tiene varios significados: palabra, estudio, tratado, discurso, sentido, entre otros. La Tanatología es una disciplina científica, es un arte y una especialidad, se ha definido como *"Disciplina encargada de encontrar sentido al proceso de la muerte"*, o como la ciencia encargada del estudio de la muerte, término que acuñó en 1901 el ruso Elie Metchnikoff, quien recibiera el premio Nóbel de medicina en 1908.

La muerte, ha despertado siempre al mismo tiempo la repugnancia y fascinación del hombre a través de los siglos. Bendecida por unos, repudiada por otros, negada por la mayoría y omnipresente en todos nosotros. El fenómeno relacionado con el término de la

¹ Este apartado se basó en J. Ferrater Mora (1994). Diccionario de filosofía, editorial Ariel como en Nicola Abbagnano (1998). Diccionario de Filosofía, Fondo de Cultura Ecómica

actividad vital no ha podido ser correctamente definido por nadie, ni en términos filosóficos, ni en términos biológicos, ni mucho menos en el plano matemático, por lo que, la mejor definición sigue siendo: *muerte es lo contrario a la vida*.

La ciencia cree saber sobre la muerte celular, hechos que no admiten duda ni disputa y que algunas frases dentro de los mismos textos de patología sorprenden cuando ponen al descubierto nuestra real ignorancia: "La transición de la vida a la muerte es igualmente difícil de precisar para la célula como para el organismo considerado globalmente". Para los filósofos la muerte no es un hecho clínico, sino un fenómeno relacionado con la pérdida de la vida, y con ello de la condición personal del ser humano.

Para los griegos la muerte se identificaba con la pérdida de funciones vitales. Tuvieron una idea funcionalista de la vida, y por ello identificaron a la muerte con la pérdida de las funciones vitales, representadas por el latido cardíaco y la respiración espontánea.

Aristóteles al comienzo del libro de la Política, dice *"que todo animal dotado de logos es humano, y que la pérdida del logos es la pérdida de la humanidad"*.

Una definición médico-legal de muerte es: la detención del proceso que preserva la integridad física del cuerpo. Es decir, la detención del proceso de la vida y como tal es una secuencia de hechos que culmina en la terminación de las funciones biofisiológicas. A este respecto, los legistas (tanto médicos como juristas) han estudiado a fondo las manifestaciones de la muerte, la que han confirmado mediante la comprobación de la desaparición de las funciones nerviosas (inmovilidad, flacidez muscular, parálisis de esfínteres, ausencia de actividades psíquicas), extinción de las funciones circulatorias (ausencia de pulso y latidos cardíacos), y extinción de las funciones respiratorias.

El concepto de muerte como finalización del tiempo y el movimiento también ha sido explorado por algunos filósofos: Wittgenstein decía: *"La muerte no es un evento de la vida. La muerte no se vive. Si por eternidad se entiende no la duración interminable del tiempo, sino la ausencia total del tiempo, entonces el que vive eternamente es aquel que vive en el presente"*.

Nietzsche escribía en el siglo pasado: *"Uno debe partir de la vida como Ulises partió de Nausica: bendiciéndola más que amándola"*.

De acuerdo a la neurología, existen tres grandes tendencias o escuelas que definen la muerte sobre las siguientes bases: aquellas que se refieren a la pérdida de atributos esenciales que identifican la naturaleza humana, aquellas que se refieren a la pérdida de la integración del organismo como un todo y las que tratan de definir la porción del encéfalo que debe dejar de funcionar irreversiblemente para que una persona pueda declararse fallecida.

Estas definiciones de muerte orientadas hacia las funciones encefálicas llegan a la conclusión de que la conciencia es la función que provee los atributos esencialmente humanos y, a la vez, integra el funcionamiento del organismo como un todo. Por tanto, se define a la muerte humana, como *la pérdida irreversible de la capacidad y del contenido de la conciencia*.

Como personas tenemos varias dimensiones: biológica, psíquica, social, histórica y espiritual. Cuando se muere, todas estas dimensiones se pierden. Pero durante este proceso se debe admitir que esto sucede cuando se tiene evidencia de daño cerebral total e irreversible, es decir, cuando las funciones intelectuales superiores perdidas privan a los seres de ser biológicamente funcionales, psíquicamente activos, socialmente útiles e históricamente en desarrollo.

Por lo tanto, para efectos prácticos, aunque todavía no se puede definir la muerte en forma matemáticamente exacta, la pérdida irrecatable del cerebro de un ser humano es el punto de no retorno, el límite médico que separa al vivo del muerto y el momento biológico en el que cualquier tratamiento solamente servirá para prolongar una agonía innecesaria, y la extensión variable de un dolor familiar. Por otra parte, en términos demográficos fallecimiento pierde su sentido etimológico de muerte natural, por lo que fallecimiento es sinónimo de muerto², por lo que en esta tesis sólo rescataremos la estadística de muertes reportadas como un suceso y trabajaremos con el número de eventos de este tipo.

1.2 Determinantes de la mortalidad

El fenómeno de la mortalidad en una población está directamente ligado a la frecuencia de la morbilidad (enfermedad y dolencias de la población) y de la probabilidad de morir de

² Naciones Unidas (1974). Diccionario Demográfico Plurilingüe, Volumen Español, USA.

una persona enferma. La muerte puede ser causada por enfermedades transmisibles, por procesos degenerativos o por causas directamente atribuibles al contexto social, económico y cultural, aunque en realidad en la mayoría de los casos la ocurrencia del evento combina varios de los factores mencionados, por ejemplo: la edad, el sexo, la raza, la ocupación y la clase social. La incidencia de las muertes proporciona información sobre el nivel de vida y de los servicios de salud de la población. La mortalidad al igual que la fecundidad y la migración es un fenómeno que influye en gran medida en los cambios estructurales de la población.

Martínez y Bronfman (1984), en su análisis sobre mortalidad, observaron que existía la certidumbre de que el suceso muerte no se distribuía aleatoriamente en los distintos grupos de la sociedad, por lo que, condujeron gran parte de su investigación a encontrar los determinantes de la mortalidad, para así poder descubrir las particularidades de cómo ocurre. Ellos consideraron que la mortalidad es la expresión final de las condiciones en las cuales transcurre la vida de la población, y su estudio, no se puede limitar a considerarse como una variable que afecta el tamaño y la estructura de la población.

La mortalidad infantil y la esperanza de vida son los indicadores que más se utilizan para evaluar el nivel de desarrollo de una sociedad. Por lo que, "El conocimiento de la magnitud y evolución de la mortalidad en la niñez resulta de importancia prioritaria en cualquier país, ya que la incidencia relativa y las causas de los fallecimientos entre los menores, además de reflejar las condiciones generales de salud y de la situación socioeconómica de las familias a donde pertenecen los niños, manifiestan la existencia y accesibilidad de la infraestructura médica y hospitalaria, la eficacia de las campañas de vacunación y erradicación de ciertas enfermedades, las características de la vivienda, etc." (Corona, 1989).

La mortalidad infantil al igual que la mortalidad general, tiene una parte importante de su origen en fenómenos biológicos (como son la edad, sexo, y características genéticas), que están íntimamente ligados con el proceso de reproducción y puede ser favorablemente modificado, mediante el cambio de los patrones de conducta reproductiva y la participación de servicios asistenciales que mejoren drásticamente la atención del embarazo, del parto y del recién nacido³.

³ Academia Mexicana de Investigación de Demografía Médica, A. C (1988). *Temas de Planificación Familiar, Mortalidad del menor de 5 años*, México.

Frente a la necesidad de obtener un mejor análisis de los factores que influyen en la ocurrencia de la mortalidad infantil, Mosley y Chen (1984), desarrollaron un marco conceptual para investigar la sobrevivencia infantil en los países en desarrollo. El marco está integrado por variables de procesos mórbidos, biológicamente determinados por otras de tipo social que se dan en la familia y la comunidad. Los factores de riesgo se agrupan en dos grandes categorías: los determinantes próximos y los subyacentes.

1.3 Determinantes próximos

Son los mecanismos biosociales básicos que influyen directamente en los riesgos de morbilidad y mortalidad.

Mosley y Chen (1984) propusieron un marco simple para el estudio de sobrevivencia en la infancia e identificaron 14 mecanismos básicos, mismos que se agruparon en cinco categorías.

- *Factores de fertilidad materna:* Edad, paridad e intervalos intergenésicos.
- *Contaminación ambiental:* aire, alimentos, agua, piel, tierra, objetos inertes e insectos vectores.
- *Deficiencias nutricionales:* energía, proteínas, micronutrientes.
- *Lesiones:* accidentales e intencionales.
- *Control personal de la enfermedad:* medidas preventivas personales y tratamiento médico.

El marco que engloba a la fecundidad y a las variables intermedias de la sobrevivencia en la infancia, identifica nueve determinantes próximos que pueden situarse en cuatro grupos relacionados con: exposición a la concepción, lactancia, riesgos ecológicos e intervenciones deliberadas. (Cuadro 1 del anexo).

Los factores de la concepción, se refieren a la formación y estabilidad de las uniones maritales. La edad materna, al igual que la separación de las parejas y el divorcio, tienen aspectos similares sobre la fecundidad y la sobrevivencia en la infancia, el matrimonio a edades tempranas y las uniones estables aumentan la fecundidad, pero también pueden contribuir directamente a una mortalidad más elevada en la infancia y niñez si favorecen la ocurrencia de nacimientos de alto orden de paridad y en los extremos de la edad reproductiva. Las uniones inestables pueden disminuir la fecundidad, pero también pueden elevar la mortalidad a través de la desorganización social en la familia.

La disminución en la prevalencia de la lactancia materna, puede elevar tanto la fecundidad como la mortalidad, si no hay un uso adecuado de anticonceptivos y de nutrientes.

Los factores ecológicos de riesgo, en su conjunto, se refieren a los efectos adversos en el desempeño reproductivo tanto bruto como neto, ya que disminuyen la fertilidad al mismo tiempo que elevan la mortalidad. Por ejemplo, la reducción en la mortalidad, proviene de enfermedades como la gonorrea que produce esterilidad y por el paludismo que ocasiona muertes fetales, así como las lesiones del tracto reproductivo femenino, que pueden disminuir la fecundidad.

Las intervenciones directas tienen una combinación de efectos diferentes. Por ejemplo, una medida preventiva como la anticoncepción no sólo reduce la fecundidad sino que también puede mejorar la sobrevivencia infantil al afectar la edad de la madre al primer y/o último nacimiento, el intervalo entre nacimientos, y el número total de niños nacidos (Rohde y Allman, 1981).

En contraste, la atención prenatal y durante el parto, a la vez que pueden hacer disminuir la mortalidad, aumenta la fecundidad al asegurar embarazos exitosos y previene lesiones en los órganos reproductivos femeninos. Los programas curativos, por ejemplo, las terapias efectivas como la penicilina, mejoran la sobrevivencia, y en circunstancias en que las enfermedades venéreas sean muy prevalentes, puede incrementar significativamente la fecundidad. Las lesiones intencionales cubren un rango de intervenciones tales como el aborto que ocasiona la reducción de la fecundidad, el infanticidio que reduce el tamaño

de la familia y las prácticas como la circuncisión femenina, que pueden afectar tanto a la fecundidad como a la sobrevivencia infantil, (Romaniuk A, 1980).

Por otra parte, en la década de los 50, Davis y Blake identificaron 11 mecanismos, que se denominan determinantes próximos (o variables intermedias), éstas se relacionan con la exposición al coito, exposición al riesgo de concepción a través del coito y la probabilidad de una gestación exitosa que conduce al nacimiento de un producto vivo. Subsecuentemente (Bongaarts, 1982) simplificó los problemas cuantitativos asociados con el uso de este enfoque, reduciendo el número de determinantes próximos a ocho variables relacionadas con la exposición al coito, factores de control deliberado de la fertilidad marital y factores de control natural de la fecundidad marital.

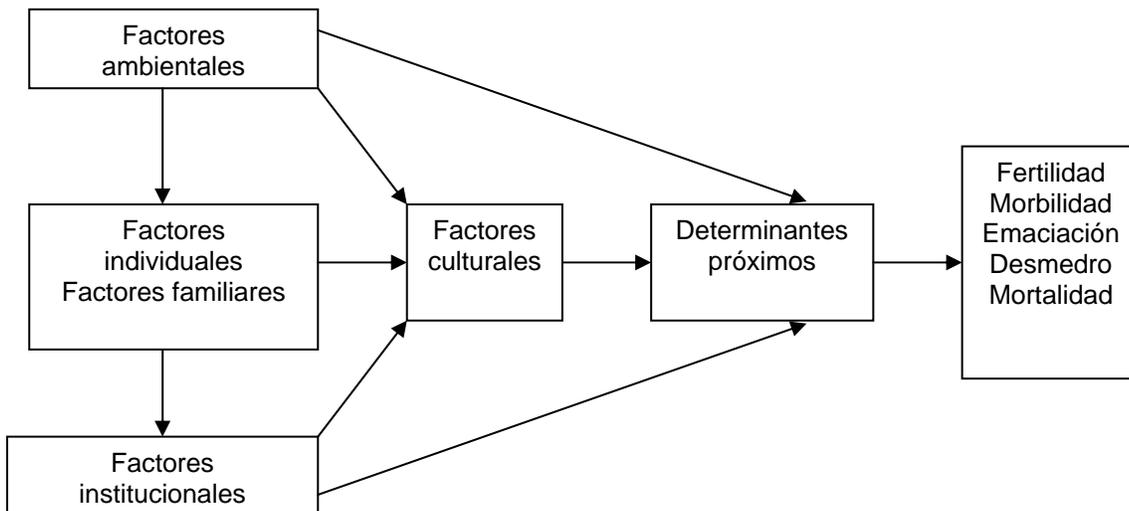
1.4 Determinantes subyacentes

Son todos los demás factores sociales y ambientales que operan indirectamente a través de los determinantes próximos, con el fin de influir en la sobrevivencia infantil. Los determinantes subyacentes de la mortalidad en la infancia y la niñez, Mosley y Chen (1988) los clasifican en cinco grupos.

- *Factores individuales:* capacidad para el trabajo de la madre, conocimiento, tiempo, salud y control sobre sus recursos.
- *Factores familiares:* capacidad para el trabajo del padre, conocimientos, salud, ingresos, bienes, servicios, ahorros, recursos y composición familiar.
- *Factores culturales:* tradiciones, normas, preferencias, gustos y valores.
- *Factores institucionales:* economía, política, infraestructura, instituciones sociales, burocracia, programas de salud y otros programas.
- *Factores ambientales:* clima, tierra, agua y topografía.

Figura 1.1

Modelo general de las interrelaciones entre los determinantes próximos y subyacentes a la fertilidad y la sobrevivencia infantil



Fuente: Determinantes biológicos y socioeconómicos de la sobrevivencia en la infancia. Henry Mosley (1988).

Los factores individuales, familiares, institucionales y ambientales, por lo general se consideran como variables independientes o exógenas, mientras que los factores culturales, que efectivamente actúan como moduladores de la conducta, pueden ser considerados como variables dependientes o endógenas. Los factores ambientales e institucionales pueden incidir directamente sobre los determinantes próximos, mientras que las variables individuales y familiares generalmente se ven restringidas por los factores culturales Mosley y Chen, (1984).

1.5 El patrón de formación familiar

La metodología que se utilizó en el presente trabajo está formulada por John Hobcraft, y para llevar a cabo el análisis se requiere la historia de embarazos, que proporciona datos de cada uno de los embarazos que tienen las mujeres, así como la condición del término del embarazo y además el tiempo que transcurrió entre cada embarazo.

La metodología consiste en combinar la edad de la madre al momento del nacimiento de sus hijos, el tiempo que transcurrió entre cada nacimiento, el orden de nacimiento y el ritmo

reproductivo. Hobcraft en la década de los 80, denominó a esta combinación de factores "Patrón de formación familiar".

El patrón de formación familiar se compone de 10 categorías, en las que se consideran a todos los nacimientos cuya edad de la madre al momento del nacimiento era menor de 35 años de edad y donde se excluyen a :

- nacimientos múltiples: dos o más productos de la concepción (embarazo);
- mortinatos o nacido muerto: producto de la concepción de 26 semanas (6 meses o más) de gestación (embarazo) y que al momento de su extracción o expulsión del vientre materno no mostró signo de vida como llanto, respiro o latido del corazón;
- abortos: interrupción inducida o espontánea del embarazo antes de que el producto de la concepción fuera viable, esto es, antes del sexto mes de gestación.

Los nacimientos se distribuyen de acuerdo a cuatro criterios:

- 1) Primeros nacimientos y nacimientos posteriores.
- 2) Nacimientos de madres jóvenes, es decir, mujeres menores de 20 años de edad y nacimientos de mujeres de entre 20 y 34 años de edad.
- 3) Los nacimientos con hermano previo o nacimientos posteriores se subdividen en nacimientos ocurridos en intervalos cortos (menor de 2 años) u ocurridos en intervalos largos (2 años o más).
- 4) En los nacimientos de mujeres entre 20 y 34 años de edad se define el ritmo reproductivo: bajo, medio o alto.

El ritmo reproductivo depende de la edad de la madre y del número de hijos que tenga en ese momento de acuerdo con los criterios establecidos en el siguiente cuadro:

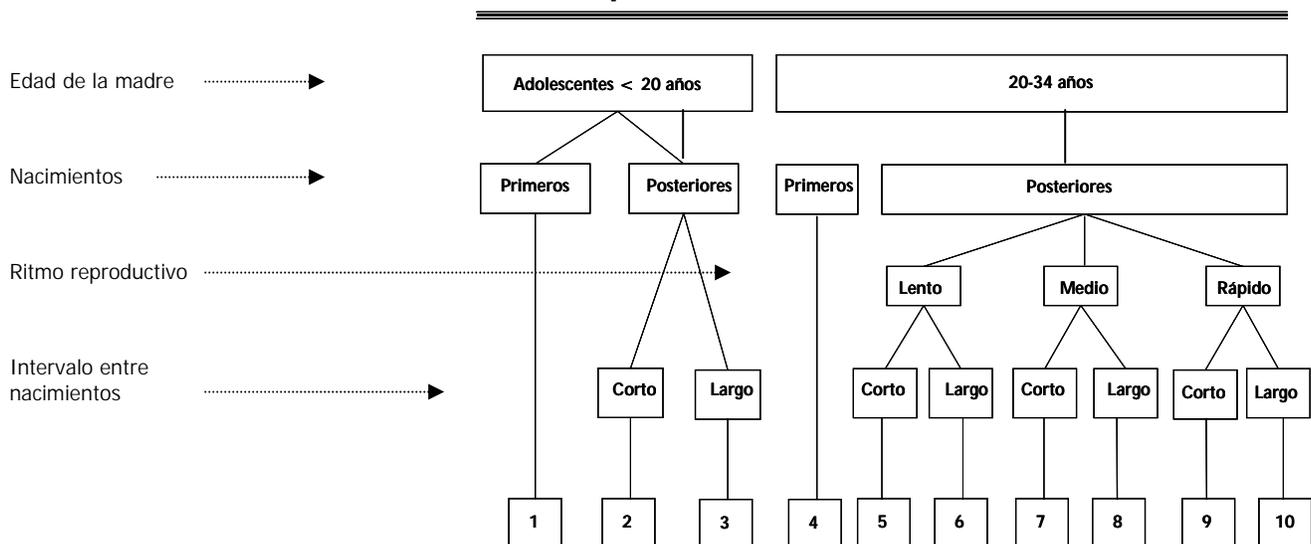
Cuadro 1.1
Ritmo reproductivo según edad de la madre
y número de hijos

Ritmo reproductivo	Grupos de Edad	Número de hijos
Bajo	20-24	2
	25-29	2-3
	30-34	2-4
Medio	20-24	3
	25-29	4-5
	30-34	5-6
Alto	20-24	4+
	25-29	6+
	30-34	7+

Fuente: Bobadilla et. al. (1990).

El siguiente esquema muestra la distribución de las 10 categorías que componen el patrón de formación familiar.

Figura 1.2
Variables para clasificar a nacidos vivos en
patrones de formación familiar



Fuente: Bobadilla et. al. (1990), pág.5.

1.6 La mortalidad infantil y el patrón de formación familiar

En 1992, John Hobcraft⁴ llevó a cabo un estudio en 25 países incluyendo a México, cuyo objetivo fue analizar la relación existente entre la fecundidad y la sobrevivencia infantil en los 80, década donde el fenómeno de la mortalidad infantil repunta y se convierte en un punto de atención prioritario, para formular políticas basadas en la disminución de muertes de niños con base en los hallazgos recientes.

Entre los resultados que obtuvo Hobcraft en la primera parte del estudio fueron: Los niños nacidos en intervalos cortos experimentan un considerable riesgo de mortalidad. Cuando hay un nacimiento posterior donde el intervalo fue menor de 18 meses, el promedio de riesgo de morir antes de los cinco años de edad es de más del doble comparado con un nacimiento posterior ocurrido en un intervalo entre dos o cuatro años.

El riesgo de mortalidad asociado con los 10 grupos del patrón de formación familiar, fue evaluado comparando la mortalidad de cualquier grupo con un "grupo de referencia". Hobcraft tomó como grupo de referencia a las mujeres con edad entre 20 y 34 años, con nacimientos posteriores y alto orden de nacimientos y con intervalos bien espaciados entre los nacimientos (categoría 10).

Entre los resultados que encontró fue que: el riesgo promedio de muerte para niños antes de cumplir los cinco años fue muy alto para nacimientos ocurridos en intervalos cortos, de mujeres jóvenes o para mujeres entre 20 y 34 años de edad, con ritmo reproductivo rápido, y la diferencia es dos y un cuarto de veces mayor con respecto al grupo de referencia.

Otro resultado fue que cuando el ritmo reproductivo era medio, el promedio de la mortalidad para los 25 países analizados fue de 82 por ciento; cuando era bajo fue de 44 por ciento. Los primeros nacimientos de madres jóvenes experimentaron un promedio de mortalidad de 47 por ciento de riesgo de mortalidad antes de cumplir los cinco años de edad. Así mismo los nacimientos de madres jóvenes con intervalos bien espaciados

⁴ Naciones Unidas publica su trabajo en *The Health Rationale for Family Planning Timing of Births and Child Survival*, 1994, New York.

experimentaron un 35 por ciento de mortalidad. Lo anterior le permitió clasificar las 10 categorías que forman el patrón de formación familiar en tres grupos:

- Bajo riesgo, compuesto de primeros nacimientos y nacimientos posteriores con intervalos bien espaciados de mujeres con edad entre los 20 y 34 años.
- Riesgo intermedio, en el que incluyó a los primeros nacimientos y nacimientos posteriores con intervalos bien espaciados, de mujeres jóvenes y ritmo reproductivo bajo y con intervalos entre nacimientos pobremente espaciados de mujeres de 20 a 34 años de edad.
- Alto riesgo compuesto por nacimientos con intervalos pobremente espaciados, de mujeres jóvenes y mujeres de 20 a 34 años de edad, con ritmo reproductivo medio y también con intervalo pobremente espaciado entre los nacimientos.

En resumen, la mortalidad de todos los nacimientos clasificados en los grupos intermedio y de alto riesgo, resultó casi tres veces mayor que el del grupo de bajo riesgo. En el análisis multivariado, Hobcraft consideró cinco modelos de regresión logística, que aplicó para cada uno de los 25 países.

- Al primer modelo se le conoce como “Cantidad de nacimientos”, con el cual se determina si el número de nacimientos en intervalos de tiempo antes del nacimiento del siguiente niño está asociado con el riesgo de muerte del niño. El número de nacimientos en los dos años previos a un nacimiento específico (pobremente espaciado) y los nacimientos ocurridos en los cuatro años precedentes se incluyen en el análisis, estos se controlan por los primeros nacimientos y los años de escolaridad de la madre. Los resultados del modelo indican que existe un fuerte efecto de los nacimientos con intervalos pobremente espaciados pero también formaliza algunos de los resultados con respecto al ritmo reproductivo.
- El segundo modelo es conocido como “biodemográfico”, el cual combina los espaciamientos entre nacimientos, la edad y la educación de la madre, el orden de nacimiento, el sexo del niño e incluye como control la sobrevivencia del hermano

previo. El cual provee un indicador burdo pero efectivo de muertes cercanas dentro de la familia.

- El tercer modelo combina diferentes factores socioeconómicos que se correlacionan con la mortalidad en la niñez, el cual incluye los años de escolaridad de la madre y del padre, la ocupación del padre y la región de residencia. Aunque este modelo no está directamente asociado con la frecuencia de nacimientos, ayuda a esclarecer la interrelación entre tales variables e indica la importancia relativa como control para un subsecuente análisis.
- El cuarto modelo se ajusta sólo a primeros nacimientos, en los cuales es necesario identificar para cada primer nacimiento la edad de la madre y el riesgo excesivo de tener nacimientos a muy temprana edad.
- El quinto modelo se aplica sólo a nacimientos posteriores (no incluye a primeros nacimientos), este modelo analiza la combinación entre: el intervalo del nacimiento anterior, la edad de la madre al momento del nacimiento y el orden de nacimiento, el estatus de sobrevivencia del nacimiento anterior y el número de muertes entre los primeros nacimientos, el sexo del niño, los años de escolaridad de la madre, la región de residencia y la tendencia en el tiempo. Este modelo por lo tanto combina diferentes grupos de variables como el estatus socioeconómico, aspectos biodemográficos de la regulación de los nacimientos, y agrupa las muertes dentro de las familias.

Otro de los trabajos que retoma la metodología de la formación de patrones de formación familiar fue realizado por José Luis Bobadilla y colaboradores en 1990.

El objetivo del trabajo fue determinar los efectos de la variación en los patrones de formación familiar de las mujeres mexicanas sobre la muerte fetal y la mortalidad en la niñez. El estudio consistió en un análisis de la Encuesta Nacional sobre Fecundidad y Salud (ENFES) levantada en 1987. El estudio incluyó a 13,216 nacimientos ocurridos en los 15 años previos a la fecha de la encuesta. Las variaciones en los patrones de formación familiar y su relación con la muerte fetal, neonatal, postneonatal, infantil y mortalidad en la niñez fueron

analizadas en forma conjunta con variables socioeconómicas aplicando un modelo de regresión logística.

Las conclusiones del estudio fueron:

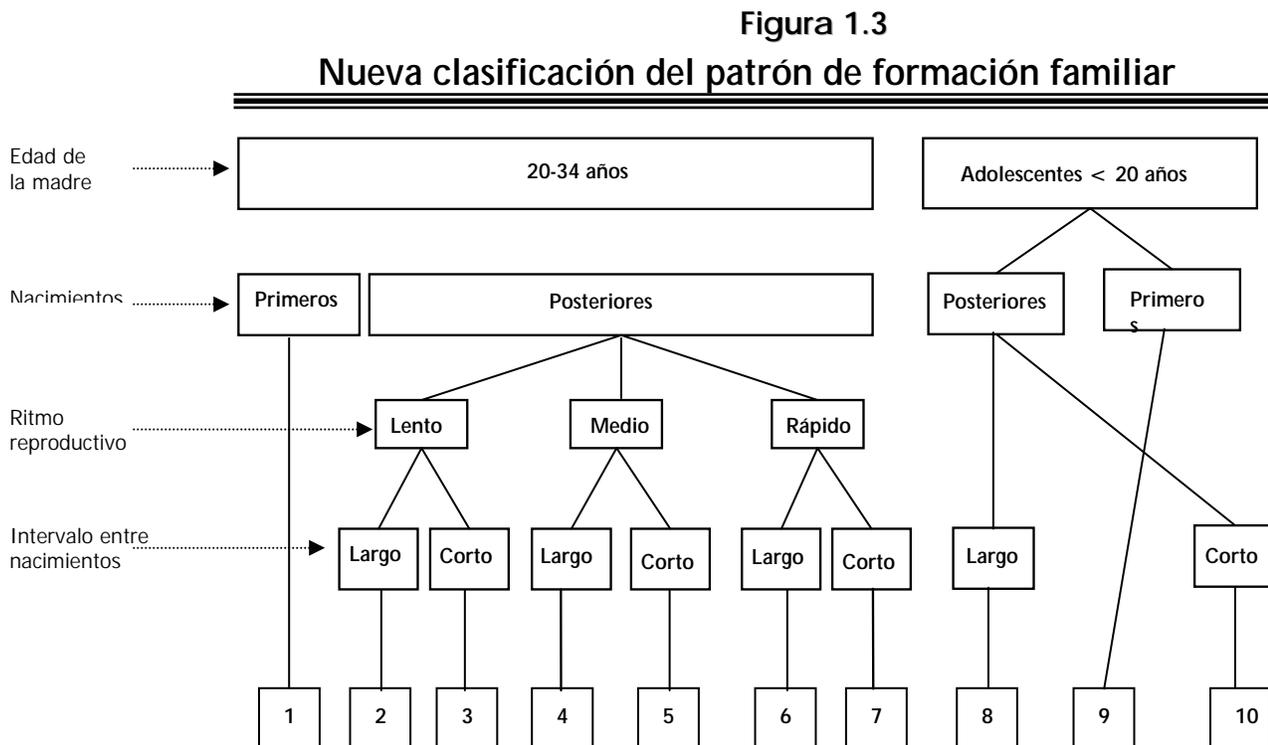
- Los grupos con alta tasa de muerte fetal están caracterizados por nacimientos con intervalos pobremente espaciados entre las mujeres de 20 y 34 años de edad, y mujeres mayores con un ritmo reproductivo medio o rápido y nacimientos de mujeres mayores de 35 años de edad.
- La mortalidad infantil entre nacimientos de mujeres jóvenes es más alta que el promedio nacional, si los intervalos entre nacimientos son pobremente espaciados. Estos mismos reflejan desventajas socioculturales para este grupo de mujeres jóvenes.
- Las tasas de mortalidad infantil de nacimientos de mujeres con edad entre 20 y 34 años, se incrementa conforme el ritmo reproductivo cambia de bajo a medio y a rápido y el intervalo entre nacimientos con hermano previo cambia de más de 24 meses a menos meses.
- Nacimientos de mujeres con edad de 35 años o más, tienen tasas de mortalidad infantil igual que la tasa nacional.
- Los efectos estimados de la mortalidad en la niñez revelan una notable similitud con la mortalidad infantil.
- Se encontró una importante relación inversa entre la mortalidad en todos los periodos estudiados y variables socioeconómicas, tales como la educación y el tamaño de la comunidad de residencia de la madre.

En resumen, estos estudios indican que existen cambios cuantitativos y cualitativos en los patrones de reproducción. Las familias mexicanas atraviesan por un proceso fuerte de transición cultural y socioeconómica. La reducción en el número de niños provee sin lugar a dudas, una importante evidencia del cambio en los valores y actitudes hacia los niños. Sin

embargo, es importante señalar que la reducción en la fecundidad no necesariamente implica cambios positivos en los patrones de formación familiar en México.

Pero también es importante señalar que en todos los estudios anteriormente mencionados, el patrón de formación familiar se analizó desde un enfoque cualitativo nominal; sin embargo, el patrón de formación familiar se puede considerar como una variable de tipo ordinal, ya que las diez categorías en que se desagrega, indican un mayor riesgo para el resultado del embarazo de la mujer, de acuerdo con los diferentes criterios con los que se definen las categorías. Por ejemplo, los embarazos, de mujeres entre las edades de 20 a 34 años y que son sus primeros embarazos tienen asociado un menor riesgo de que el producto de ese embarazo muera en el primer año de vida, que aquellas mujeres que se embarazan a una edad menor de 20 años y que el intervalo entre el embarazo anterior es corto. Por lo que, resulta importante analizar en México el efecto del patrón de formación familiar en la sobrevivencia o muerte de los niños en el primer año de vida desde una dimensión ordinal.

La figura 1.3 muestra la nueva reclasificación de las categorías.



Fuente: Diseño propio con base en Bobadilla et. al. (1990), pág.5.

CAPÍTULO II

El modelo de regresión logística

2.1 Tipos de variables

En una población de estudio las características que cambian de un elemento de la población a otro, se conocen como variables, las cuales pueden ser de diferentes tipos. La importancia de identificar los diferentes tipos de variables se debe a que el análisis estadístico depende del tipo de variable de que se trate. Los diferentes tipos se pueden encontrar se muestran en el siguiente esquema.



Se dice que una **variable es continua** cuando entre dos valores cualesquiera puede existir un número infinito de valores intermedios. Dos características importantes de estas variables es que sus valores se pueden operar algebraicamente y que existe una relación de orden entre ellos. Ejemplos de este tipo de variables son: la edad medida en años y fracción de año. Un posible valor de esta variable es 15.4 que es mayor que 12 y que 15.3, pero menor que 15.6; además la suma de las edades 12 y 16 da como resultado 28 años que es un valor válido de la variable, otro ejemplo es el salario medido en pesos y centavos. Una persona que gana 570.32 pesos al mes, gana más que alguien cuyo salario es de 325.62 pesos al mes, pero menos que el que gana 1,225.2 pesos. Además alguien puede ganar un salario de $570.32+325.62+1,225.2=2,121.14$ pesos al mes, el cual se obtiene sumando los tres salarios mencionados anteriormente.

Las **variables discretas** son aquellas entre cuyos dos valores cualesquiera consecutivos no existe otro valor. También se caracterizan por una relación de orden y porque al operar algebraicamente algunos de sus valores (no todos), se obtiene nuevamente un valor válido de la variable. Por lo general este tipo de variables surgen de efectuar operaciones de conteo. Ejemplos de estas variables son el número de hijos de una mujer, cuyos valores válidos pueden ser 0, 1, 2, o 3 hijos, pero un valor de 2.5 no es válido, ya que no se puede tener un pedazo de hijo. El número de empleos es otro ejemplo, cuyos valores 0, 1, 2, 3, o 4 pueden ser válidos, pero no se pueden tener 3.5 empleos, etc.

Las **variables de tipo ordinal** son aquellas cuyos valores establecen un orden, es decir, las respuestas obtenidas tienen implícito un valor que sirve para definir cuál valor es mayor y cual es menor, pero no se puede definir una distancia entre ellos ni determinar cuántas unidades separan a los valores, solamente se sabe que, uno es mayor que otro pero no se sabe en cuántas unidades es mayor. Por ejemplo en la variable escolaridad, la clasificación secundaria se sabe que es mayor que la clasificación primaria, pero la secundaria no se obtiene sumando uno o dos niveles de escolaridad a la primaria. Otro ejemplo, puede ser la efectividad de un método anticonceptivo, ya que se sabe que el dispositivo intrauterino es más efectivo que las pastillas, pero la efectividad del dispositivo no se obtiene sumando 1, 2 o 3 niveles de efectividad a las pastillas.

Las **variables nominales** o categóricas, se caracterizan porque los valores que toman no pueden ser sometidos a un criterio de orden y no son operacionales, las respuestas de este

tipo de variables se pueden identificar asignándoles números o letras o cualquier otro identificador, pero las etiquetas por si solas no significan nada. Algunos ejemplos de este tipo de variables son: el sexo con dos categorías hombre y mujer; el estado civil con cuatro categorías soltera, casada, viuda y divorciada; y la condición de trabajo con las categorías si trabaja y no trabaja.

2.2 Regresión logística simple¹

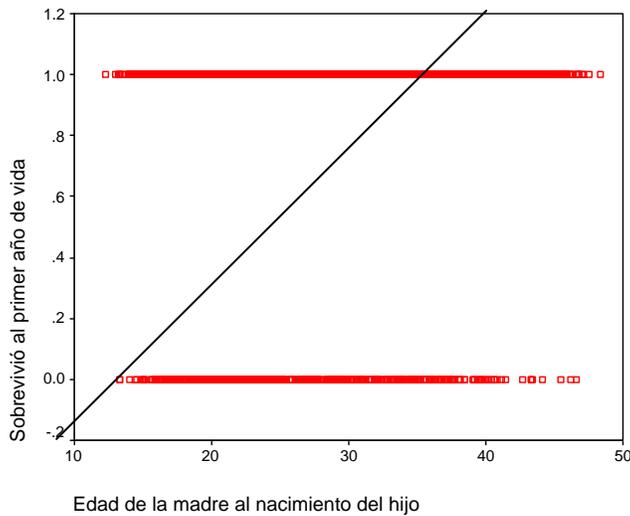
La regresión logística simple se utiliza para explicar el comportamiento de una variable dependiente binaria de tipo nominal en función de una variable continua, discreta, ordinal o categórica.

Como se trata de una variable dependiente binaria nominal, entonces sus valores no pueden crecer ni decrecer porque no existe una relación de orden entre sus valores. Sin embargo, la probabilidad de observar alguno de sus valores si tiene una relación de orden, por lo que la regresión logística simple, busca explicar el comportamiento de la probabilidad de uno de los valores en función de otra variable.

Por ejemplo, si se desea determinar la sobrevivencia del niño al primer año de vida en función de la edad de su madre, entonces se tiene que la variable independiente es continua y la dependiente dicotómica, cuyos posibles valores son, que sobreviva el niño o que no sobreviva, y que la relación de interés buscada es determinar si la probabilidad de sobrevivencia se incrementa o decrecienta en tanto que la edad se incrementa también. Un primer intento para explicar esta situación, es utilizar una línea recta al igual que en el modelo de regresión lineal simple (Figura 2.1).

¹ Este capítulo se basó principalmente en Hosmer D. y Lemeshow S. (1989). Applied Logistic Regression, Massachussets, USA, desarrollandose en detalle las fórmulas.

Figura 2.1
Probabilidad de sobrevivir al primer año de vida según edad de la madre al nacimiento del hijo



Fuente: Hosmer D. y Lemeshw Applied Logistic Regresión.

El uso de estas funciones no es del todo adecuada, ya que siguiendo el comportamiento de la línea recta, la probabilidad puede asumir valores mayores que uno y menores que cero, lo cual no es posible al estudiar probabilidades (Figura 2.1). El problema se resuelve utilizando una función que solamente puede tomar valores entre cero y uno. Dicha función se conoce como función logística y su forma matemática es:

$$P(Y = 1) = \frac{1}{1 + e^{a+bx}} \quad (2.1)$$

Donde:

$P(Y = 1)$ es la probabilidad de sobrevivencia,

a y b son parámetros desconocidos, y

x es la edad de la madre al nacimiento del hijo.

Cuando la variable independiente es también nominal con sólo dos posibles valores, y la variable dependiente sigue siendo nominal binaria, una primera opción es analizar la asociación de las variables mediante una tabla de 2 x 2 (Reynolds, 1977). Un ejemplo de esta situación puede ser, aquella en la que se desea saber si el sexo del niño influye para que éste sobreviva al primer año de vida. En este caso los posibles valores de la variable independiente es que sea hombre o mujer y de la dependiente que viva o muera al primer

año de vida. En el Cuadro 2.1 se presenta un análisis de este tipo de variables, con base en el ENADID de 1997.

Cuadro 2.1
México. Condición de sobrevivencia al primer año de vida por sexo

Condición de sobrevivencia al primer año de vida	Sexo del niño	
	Hombre	Mujer
No sobreviva	286	254
Sobreviva	11,210	10,795
Total	11,496	11,049

Fuente: INEGI. Encuesta Nacional de la Dinámica Demográfica, 1997.
 Nota: Se estiman los nacimientos de 1994 a 1996.

De acuerdo con los datos del Cuadro 2.1, se puede observar que en México nacen más niños que niñas, aproximadamente 104 por cada 100 niñas; sin embargo, también se observa que mueren más niños que niñas, casi 112 niños por cada 100 niñas. Con estos resultados se podría concluir que las niñas tienen menos riesgo de morir antes de cumplir el primer año de vida.

Otra medida de asociación en una tabla de 2 x 2 es el cociente de momios, llamado también cociente del producto cruzado, y que se define como:

$$\psi = \frac{\frac{P_{11}}{P_{21}}}{\frac{P_{12}}{P_{22}}} = \frac{P_{11}P_{22}}{P_{21}P_{12}} \quad (2.2)$$

Donde:

P_{11} es la frecuencia del renglón 1 columna 1, en una tabla de 2 x 2,

P_{12} es la frecuencia del renglón 1 columna 2, en una tabla de 2 x 2,

P_{21} es la frecuencia del renglón 2 columna 1, en una tabla de 2 x 2,

P_{22} es la frecuencia del renglón 2 columna 2, en una tabla de 2 x 2,

ψ es el cociente de momios.

El cociente puede tomar valores entre cero y más infinito, cuando toma el valor de uno indica independencia de variables, es decir, que no existe asociación, pero si toma valores

menores que uno indica que existe una asociación negativa y si es mayor que uno, la asociación entre las dos variables es positiva. Aplicando la ecuación a los datos del Cuadro 2.1 se tiene que:

$$\psi = \frac{(286)(10,795)}{(11,210)(254)} = 1.0843$$

El resultado muestra que existe una asociación positiva entre las variables, es decir que, los niños tienen 1.08 mayor riesgo de morir antes de cumplir su primer año de vida que las niñas.

Otra opción para analizar este tipo de datos es aplicar el modelo de regresión logística simple, mediante el cual se puede calcular el cociente de momios cuando la variable independiente es dicotómica. Para ello es necesario encontrar una expresión equivalente a la ecuación (2.1), el desarrollo es el siguiente:

$$P(Y = 1) = \frac{1}{1 + e^{a+bx}} = \frac{1}{1 + \frac{1}{e^{-(a+bx)}}} \quad (2.3)$$

si multiplicamos la ecuación 2.3 por $\frac{e^{-(a+bx)}}{e^{-(a+bx)}}$ entonces,

$$P(Y = 1) = \frac{1}{1 + \frac{1}{e^{-(a+bx)}}} \frac{e^{-(a+bx)}}{e^{-(a+bx)}} = \frac{e^{-(a+bx)}}{e^{-(a+bx)} + 1} = \frac{e^{-a-bx}}{1 + e^{-a-bx}} \quad (2.4)$$

si $a_0 = -a$ y $b_1 = -b$, entonces se tiene que:

$$P(Y = 1) = \frac{e^{a_0+b_1x}}{1 + e^{a_0+b_1x}} \quad (2.5)$$

Ahora como $P(Y = 0)$ es igual a $1 - P(Y = 1)$ entonces,

$$P(Y = 0) = 1 - \frac{e^{a_0+b_1x}}{1 + e^{a_0+b_1x}} = \frac{1 + e^{a_0+b_1x} - e^{a_0+b_1x}}{1 + e^{a_0+b_1x}} = \frac{1}{1 + e^{a_0+b_1x}} \quad (2.6)$$

Como x es una variable dicotómica el modelo logístico puede tomar sólo dos valores, cuando $x=1$ y cuando $x=0$. Se acostumbra denotar a $P(Y=1)$ como $\pi(x) = E(Y|x)$ y que denota a la probabilidad como una función de la variable x . De manera análoga se tiene que $P(Y=0) = 1 - \pi(x) = 1 - E(Y|x)$ que también denota a la probabilidad como una función de x . Se muestran los valores que asume el modelo de regresión logística (2.5) o (2.6) en los diferentes casos en el Cuadro 2.2. A la cual se le aplica la fórmula del cociente de momios (Ecuación 2).

Cuadro 2.2
Expresiones del modelo logístico simple según variable dependiente e independiente

Variable dependiente Y	Variable independiente X	
	$x=1$	$x=0$
$y=1$	$\pi(1) = \frac{e^{a_0 + b_1}}{1 + e^{a_0 + b_1}}$	$\pi(0) = \frac{e^{a_0}}{1 + e^{a_0}}$
$y=0$	$1 - \pi(1) = \frac{1}{1 + e^{a_0 + b_1}}$	$1 - \pi(0) = \frac{1}{1 + e^{a_0}}$
Total	1.0	1.0

El cociente de momios calculado con el cuadro 2.2 y con base en la fórmula (2.2) queda como:

$$\psi = \frac{\pi(1)[1 - \pi(0)]}{[1 - \pi(1)]\pi(0)}$$

Usando las expresiones que se presentan en el Cuadro 2.2 en el modelo de regresión logística se obtiene lo siguiente:

$$\psi = \frac{\left(\frac{e^{a_0+b_1}}{1+e^{a_0+b_1}}\right)\left(\frac{1}{1+e^{a_0}}\right)}{\left(\frac{1}{e^{a_0+b_1}}\right)\left(\frac{e^{a_0}}{1+e^{a_0}}\right)} = \frac{\frac{e^{a_0+b_1}}{1+e^{a_0}+e^{a_0+b_1}+e^{2a_0+b_1}}}{\frac{e^{a_0}}{1+e^{a_0}+e^{a_0+b_1}+e^{2a_0+b_1}}} = \frac{e^{a_0+b_1}}{e^{a_0}} = \frac{e^{a_0}e^{b_1}}{e^{a_0}} = e^{b_1}$$

Por lo que el modelo de regresión logística para una variable independiente dicotómica calcula el cociente de momios como e^{b_1} , donde b_1 es el coeficiente de la variable independiente x .

Al aplicar el modelo de regresión logística a los datos de la ENADID de 1997 para encontrar la asociación que existe entre las variables, es decir, para determinar si tiene mayor riesgo de no sobrevivir al primer año de vida un niño o una niña, se obtuvieron los siguientes resultados: El número total de casos procesados por el modelo fueron 22,546 casos, la variable dependiente se denotó como 1 para los que sobrevivieron al primer año de vida y 0 para los que fallecieron antes de cumplir un año, en la variable independiente se denotó como 1 a hombres y 0 a mujeres. El modelo calculó una constante (a_0) igual a 3.6686 y un coeficiente (b_1) de 0.0809, lo que da un cociente de momios $e^{b_1} = 1.0843$. Con lo que se puede corroborar que es el mismo resultado que se obtuvo cuando se analizó la asociación entre las mismas variables con una tabla de 2×2 .

También se puede analizar el caso cuando la variable independiente sea nominal con más de dos categorías, por ejemplo si se desea saber en que lugar de residencia ocurre en mayor o menor medida la mortalidad de menores de un año. El lugar de residencia tiene tres categorías: rural, urbana o metropolitana.

Los datos muestrales de la ENADID de 1997 sobre la sobrevivencia o no de los niños(as) dependiendo del lugar de residencia se presentan en el Cuadro 2.3.

Cuadro 2.3
México. Condición de sobrevivencia al primer año
de vida por lugar de residencia

Condición de sobrevivencia	Rural	Urbana	Metropolitana	Total
Si	7,516	6,150	8,339	22,005
No	265	138	137	540
Total	7,781	6,288	8,476	22,545
Cociente de momios	1.0	1.5713	2.1461	

Fuente: INEGI. Encuesta Nacional de la Dinámica Demográfica, 1997.

Nota: Se estiman los nacimientos de 1994 a 1996.

En este caso también se puede aplicar el cociente de momios, sólo que se toma una de las categorías como referencia y el resultado que se obtiene se lee en relación a esta categoría en términos de que tan probable² con respecto a dicha categoría ocurre el evento. Para este ejercicio se fijó la categoría rural.

De esta manera, el cociente de momios de los niños que residen en una zona urbana con respecto a los niños que residen en una zona rural de acuerdo con los datos del Cuadro 2.3 es:

$$\psi = \frac{(6,150 \times 265)}{(7,516 \times 138)} = 1.5713$$

Mientras que para los residentes en zonas metropolitanas con respecto a los residentes en zonas rurales es de:

$$\psi = \frac{(8,339 \times 265)}{(7,516 \times 137)} = 2.1461$$

Los cocientes de momios indican que en las zonas urbanas la propensión de sobrevivir de los niños a su primer año de vida es 1.6 veces mayor que en las zonas rurales y en las zonas metropolitanas la propensión de sobrevivir es un poco más del doble.

² El cociente de momios indica el número de veces que puede ocurrir un evento con respecto a la ocurrencia de otro, pero esto no es exactamente una probabilidad ya que no va a tomar valores entre 0 y 1, por lo que de ahora en adelante me referiré a este resultado como propensión.

Para aplicar el modelo de regresión logística en esta situación, es decir, en el caso de variables nominales con más de dos categorías, se crean tantas variables dicotómicas como el número de categorías menos una que tiene la variable nominal. A estas nuevas variables se les denomina en la literatura estadística con el nombre de "dummy", traducidas al español como variables internas, indicadoras o variables de diseño (Moliner, 2001). El modelo calcula el cociente de momios de cada categoría con respecto a la que se asigne como variable de referencia. El Cuadro 2.4 muestra los valores que asumen las variables de diseño, en el caso específico de la variable lugar de residencia:

Cuadro 2.4
Variabes dummy de lugar de residencia de la mujer

Lugar de residencia	Variables de diseño	
	D ₁	D ₂
Rural	0	0
Urbana	1	0
Metropolitana	0	1

En la aplicación del modelo logístico se introdujeron las variables diseño como dos variables independientes y la variable dependiente se codificó como 1 para los niños que sobrevivieron al primer año de vida y 0 a los que fallecieron antes de cumplir un año. El número total de casos procesados por el modelo fueron 22,546 casos. El modelo calculó una constante (a_0) igual a 3.3451 y un coeficiente (b_1) de 0.4519 y (b_2) de 0.7637 correspondientes a las variables diseño urbano y metropolitano respectivamente, lo que da un cociente de momios $e^{b_1} = 1.5713$ y $e^{b_2} = 2.1461$ corroborando así, los datos que se obtuvieron anteriormente.

Es importante aclarar que cuando la variable independiente es continua también se puede interpretar en términos del cociente de momios, el desarrollo es el siguiente:

Al modelo de regresión logística se le aplica la transformación logit, la cual se define como:

$$g(x) = \ln\left(\frac{P(y)}{1 - P(y)}\right), \text{ entonces sustituyendo los valores de la ecuación 2.5 se tiene que:}$$

$$g(x) = \ln \left[\frac{\frac{e^{a_0+b_1x}}{1+e^{a_0+b_1x}}}{1-\frac{e^{a_0+b_1x}}{1+e^{a_0+b_1x}}} \right] \Rightarrow g(x) = \ln \left[\frac{\frac{e^{a_0+b_1x}}{1+e^{a_0+b_1x}}}{\frac{1+e^{a_0+b_1x}-e^{a_0+b_1x}}{1+e^{a_0+b_1x}}} \right] \Rightarrow g(x) = \ln \left[\frac{e^{a_0+b_1x}}{1+e^{a_0+b_1x}} \right]$$

$$\Rightarrow g(x) = \ln \left[\frac{e^{a_0+b_1x}(1+e^{a_0+b_1x})}{1+e^{a_0+b_1x}} \right] \Rightarrow g(x) = \ln(e^{a_0+b_1x}) \Rightarrow$$

$$g(x) = a_0 + b_1x \quad (2.7)$$

Donde:

a_0 y b_1 son parámetros desconocidos, y

x es la variable independiente.

En el modelo (2.7) la proporción de cambio en una unidad en la variable x va a estar dado por el coeficiente b_1 , esto es, $b_1 = g(x+1) - g(x)$. Generalizando a c unidades de cambio en x la ecuación será $g(x+c) - g(x) = cb_1$ y para dar una interpretación en términos del cociente de momios se le aplica la exponencial a la diferencia logit, es decir $\psi(c) = \psi(x+c, x) = \exp(cb_1)$.

Retomando el ejemplo en el que se desea determinar la sobrevivencia del niño al primer año de vida en función de la edad de la madre al momento del nacimiento del mismo, se puede aplicar el modelo codificando la variable dependiente como 1 para los que sobrevivieron al primer año de vida y 0 para los que fallecieron antes de cumplir un año y considerando como variable independiente la edad de la mujer al momento del nacimiento del niño. Se procesaron 22,546 casos con los cuales se obtuvo una $b_1 = -0.0008$, al aplicar la exponencial en b_1 , se estima el cociente de momios en 0.9992, lo que indica que la posibilidad de sobrevivencia del niño disminuye aproximadamente en una unidad cada vez que aumenta en un año la edad de la madre al momento del nacimiento del hijo.

Cuando la variable independiente es de tipo ordinal, como en el caso en el que se desea conocer qué tanto influye el orden de nacimiento para que el niño logre sobrevivir a su

primer año de vida, el análisis es similar al caso continuo. La variable independiente en este caso tiene los valores de primero, segundo, hasta el sexto o más en orden de nacimiento, la variable dependiente es 1 para los que sobreviven al primer año de vida y cero para los que fallecen antes de cumplir el primer año. En el proceso se introdujeron 22,546 casos, se estimó una $b_1 = -0.1290$, en forma similar a una variable continua se le aplica la exponencial para obtener el cociente de momios de 0.8790, lo que indica que la posibilidad de sobrevivencia de un niño en su primer año de vida disminuye conforme aumenta el orden de nacimiento al que le corresponda nacer.

2.3 Regresión logística múltiple

El modelo logístico múltiple es aquel que incluye al mismo tiempo más de una variable independiente. Su ecuación tiene la siguiente forma:

$$P(Y = 1) = \frac{e^{a_0 + b_1 X_1 + b_2 X_2 + \dots + b_k X_k}}{1 + e^{a_0 + b_1 X_1 + b_2 X_2 + \dots + b_k X_k}} \quad (2.8)$$

Donde:

$P(Y=1)$ representa la probabilidad de ocurrencia de un evento de interés,

$a_0, b_1, b_2, b_3, \dots, b_k$ son parámetros desconocidos, y

$X_1, X_2, X_3, \dots, X_k$ son las variables independientes.

Las variables independientes X_1, X_2, \dots, X_k pueden ser todas de un solo tipo o bien combinaciones de diferentes tipos de variables.

2.4 Estimación de parámetros

El método de máxima verosimilitud consiste en examinar los valores de una muestra dada y escoger como estimadores de los parámetros desconocidos aquellos para los cuales la probabilidad de obtener los valores de la muestra es un máximo.

Si x_1, x_2, \dots, x_n son los valores de una muestra aleatoria de una población con el parámetro θ , entonces la función de verosimilitud está dada por $L(\theta) = f(x_1, x_2, \dots, x_n; \theta)$ para los valores de θ dentro de un dominio dado. En este caso $f(x_1, x_2, \dots, x_n; \theta)$ es el valor de la distribución de probabilidad conjunta o de la densidad de probabilidad conjunta de las variables aleatorias X_1, X_2, \dots, X_n en $X_1 = x_1, X_2 = x_2, \dots, X_n = x_n$.

De esta manera, si $\hat{\theta}$ es una función de las observaciones x_1, x_2, \dots, x_n y es tal que el valor de θ maximiza $L(\theta)$, entonces $\hat{\theta}$ es el estimador de máxima verosimilitud de θ .

Ejemplo 1. Se desea saber cuál es el estimador de máxima verosimilitud del parámetro θ de la distribución binomial, dado que se tienen k "éxitos" en n intentos. La distribución binomial para estimar k "éxitos" en n intentos está dada por:

$$b(x; n, \theta) = \binom{n}{k} \theta^k (1-\theta)^{n-k} \quad \text{para } k = 0, 1, 2, \dots, n$$

Entonces la función de verosimilitud está dada por

$$L(\theta) = \prod_{i=1}^n \binom{n}{x_i} \theta^{x_i} (1-\theta)^{1-x_i}$$

$$L(\theta) = \prod_{i=1}^n \binom{n}{x_i} \left[\theta^{\sum x_i} (1-\theta)^{n-\sum x_i} \right]$$

Aplicando el logaritmo natural a la función se tiene que

$$\Rightarrow \ln L(\theta) = \ln \left[\prod_{i=1}^n \binom{n}{x_i} \right] + \ln \theta^{\sum x_i} + \ln (1-\theta)^{n-\sum x_i}$$

$$\ln L(\theta) = \ln \left[\prod_{i=1}^n \binom{n}{x_i} \right] + \sum x_i \ln \theta + (n - \sum x_i) \ln(1-\theta) \quad (2.9)$$

Al obtener la primera derivada de (2.9) e igualar a cero se obtiene el máximo de la función de verosimilitud.

$$\frac{\partial \ln(\theta)}{\partial \theta} = \sum x_i \frac{1}{\theta} - (n - \sum x_i) \frac{1}{1-\theta}$$

$$\begin{aligned} \Rightarrow \frac{\partial \ln(\theta)}{\partial \theta} &= \frac{\sum x_i}{\theta} - \frac{n - \sum x_i}{1 - \theta} \\ \Rightarrow \text{si } \frac{\sum x_i}{\theta} - \frac{n - \sum x_i}{1 - \theta} &= 0 \\ \Rightarrow \frac{\sum x_i}{\theta} &= \frac{n - \sum x_i}{1 - \theta} \\ \Rightarrow \frac{1 - \theta}{\theta} &= \frac{n - \sum x_i}{\sum x_i} \\ \Rightarrow \frac{1}{\theta} - 1 &= \frac{n}{\sum x_i} - 1 \\ \Rightarrow \frac{1}{\theta} &= \frac{n}{\sum x_i} \end{aligned}$$

$\Rightarrow \theta = \frac{\sum x_i}{n}$ que es el estimador de máxima verosimilitud del parámetro binomial θ y se

denota como $\hat{\theta} = \frac{\sum x_i}{n}$.

Ejemplo 2. Sean X_1, X_2, \dots, X_n componentes de una muestra aleatoria de tamaño n de una población normal con media μ y σ^2 , encontrar la estimación conjunta de máxima verosimilitud de la media.

La función de verosimilitud está dada por $L(\mu, \sigma^2) = \prod_{i=1}^n n(x_i; \mu, \sigma)$

$$\Rightarrow L(\mu) = \left(\frac{1}{\sigma \sqrt{2\pi}} \right)^n e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2}$$

Aplicando el \ln a la función se tiene que

$$\begin{aligned} \ln L(\mu) &= \ln \left(\frac{1}{\sigma \sqrt{2\pi}} \right)^n + \ln \left(e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2} \right) \\ \Rightarrow \ln L(\mu) &= n \ln \left(\frac{1}{\sigma \sqrt{2\pi}} \right) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \end{aligned}$$

$$\Rightarrow \ln L(\mu) = n \ln \left(\frac{1}{\sigma \sqrt{2\pi}} \right) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i^2 - 2x_i\mu + \mu^2)$$

$$\Rightarrow \ln L(\mu) = n \ln \left(\frac{1}{\sigma \sqrt{2\pi}} \right) - \frac{\sum_{i=1}^n x_i^2}{2\sigma^2} + \frac{\sum_{i=1}^n x_i \mu}{\sigma^2} - \frac{n\mu^2}{2\sigma^2}$$

Se obtiene ahora la derivada parcial con respecto a μ

$$\frac{\partial L(\mu)}{\partial \mu} = \frac{\sum_{i=1}^n x_i}{\sigma^2} - \frac{2n\mu}{2\sigma^2} \Rightarrow \frac{\sum_{i=1}^n x_i}{\sigma^2} - \frac{n\mu}{\sigma^2}$$

Ahora se iguala a cero la expresión anterior para obtener el máximo.

$$\frac{\sum_{i=1}^n x_i}{\sigma^2} - \frac{n\mu}{\sigma^2} = 0$$

$$\Rightarrow \frac{\sum_{i=1}^n x_i}{\sigma^2} = \frac{n\mu}{\sigma^2} \Rightarrow \frac{\sigma^2 \sum_{i=1}^n x_i}{\sigma^2} = n\mu \Rightarrow \sum_{i=1}^n x_i = n\mu \Rightarrow \mu = \frac{1}{n} \sum_{i=1}^n x_i$$

por lo que $\hat{\mu} = \bar{X}$ que es el estimador de máxima verosimilitud.

Con el fin de aplicar el método de máxima verosimilitud en el modelo de regresión logística, es necesario construir primero la función de verosimilitud. Esta función expresa la probabilidad de los datos observados como una función de los parámetros desconocidos.

Si Y es codificado como cero o uno, entonces la siguiente expresión

$$\pi(x) = \frac{e^{b_0 + b_1 x}}{1 + e^{b_0 + b_1 x}}$$

proporciona la probabilidad condicional de que Y sea igual a 1 dado x, se denota como $P(Y=1 | x)$, si $P(Y=0 | x)$ entonces la función es $1 - \pi(x)$. De esta manera para las parejas (x_i, y_i) , donde $y_i=1$ es la contribución a la función de verosimilitud es $\pi(x_i)$ y para aquellos pares

donde $y_i=0$, la contribución a la función de verosimilitud es $1-\pi(x_i)$. Una manera de expresar la contribución a la función de verosimilitud de la pareja (x_i, y_i) es con el término:

$$\zeta(x_i) = \pi(x_i)^{y_i} [1 - \pi(x_i)^{1-y_i}] \quad (2.10)$$

Como se supone que las observaciones son independientes, la función de verosimilitud se obtiene como el producto de los términos de 2.10, es decir:

$$L(\beta) = \prod_{i=1}^n \zeta(x_i) \quad (2.11)$$

Aplicando el ln a la función se tiene que

$$\begin{aligned} L(b) &= \ln(L(b)) = \sum_{i=1}^n \{y_i \ln[\pi(x_i)] + (1 - y_i) \ln[1 - \pi(x_i)]\} \quad (2.12) \\ \ln(L(b)) &= \sum_{i=1}^n \left\{ y_i \ln \left[\frac{e^{b_0 + b_1 x_i}}{1 + e^{b_0 + b_1 x_i}} \right] + (1 - y_i) \ln \left[1 - \frac{e^{b_0 + b_1 x_i}}{1 + e^{b_0 + b_1 x_i}} \right] \right\} \\ \ln(L(b)) &= \sum_{i=1}^n \left\{ y_i \ln(e^{b_0 + b_1 x_i}) - \ln(1 + e^{b_0 + b_1 x_i}) + (1 - y_i) \ln \left[\frac{1}{1 + e^{b_0 + b_1 x_i}} \right] \right\} \\ \ln(L(b)) &= \sum_{i=1}^n \{y_i (b_0 + b_1 x_i) - \ln(1 + e^{b_0 + b_1 x_i}) + (1 - y_i) [-\ln(1 + e^{b_0 + b_1 x_i})]\} \quad (2.13) \end{aligned}$$

Obteniendo la derivada parcial de 2.13 con respecto a b_0 e igualando a cero.

$$\begin{aligned} \frac{\delta(L(b))}{\delta b_0} &= \sum_{i=1}^n \left\{ y_i \left[1 - \frac{1}{1 + e^{b_0 + b_1 x_i}} e^{b_0 + b_1 x_i} \right] + (1 - y_i) \left[-\frac{1}{1 + e^{b_0 + b_1 x_i}} e^{b_0 + b_1 x_i} \right] \right\} \\ \frac{\delta(L(b))}{\delta b_0} &= \sum_{i=1}^n \{y_i [1 - \pi(x_i)] + (1 - y_i) [-\pi(x_i)]\} \\ \frac{\delta(L(b))}{\delta b_0} &= \sum_{i=1}^n \{y_i - y_i \pi(x_i) - \pi(x_i) + y_i \pi(x_i)\} \\ &= \sum_{i=1}^n [y_i - \pi(x_i)] = 0 \quad (2.14) \end{aligned}$$

Ahora se obtiene la derivada parcial de 2.13 con respecto a b_1 y se iguala a cero.

$$\frac{\delta(L(b))}{\delta b_1} = \sum_{i=1}^n \left\{ y_i \left[x_i - \left(\frac{1}{1 + e^{b_0 + b_1 x_i}} x_i e^{b_0 + b_1 x_i} \right) \right] + (1 - y_i) \left[-\frac{1}{1 + e^{b_0 + b_1 x_i}} x_i e^{b_0 + b_1 x_i} \right] \right\}$$

$$\begin{aligned}\frac{\delta(L(b))}{\delta b_1} &= \sum_{i=1}^n \{y_i[x_i - (x_i\pi(x_i))] + (1 - y_i)[-x_i\pi(x_i)]\} \\ \frac{\delta(L(b))}{\delta b_1} &= \sum_{i=1}^n \{y_i x_i - y_i x_i \pi(x_i) - x_i \pi(x_i) + y_i x_i \pi(x_i)\} \\ \sum_{i=1}^n x_i [y_i - \pi(x_i)] &= 0\end{aligned}\quad (2.15)$$

“Las ecuaciones 2.14 y 2.15 se conocen como las ecuaciones de verosimilitud y no son lineales en b_0 y b_1 , por lo que se requiere de un método especial para su solución. Estos métodos son iterativos y han sido programados en los paquetes estadísticos para regresión logística” (Hosmer y Lemeshow, 1989, pág 10).

El método de estimación utilizado en el caso multivariado es el mismo que en el caso univariado (máxima verosimilitud). La función de verosimilitud es idéntica a la de 2.11 con la única diferencia de que aparece $\pi(\bar{x})$ donde \bar{x} es un vector tal que:

$$\pi(x) = \frac{e^{b_0 + b_1 X_1 + b_2 X_2 + \dots + b_p X_p}}{1 + e^{b_0 + b_1 X_1 + b_2 X_2 + \dots + b_p X_p}}$$

En este caso se obtienen $p+1$ ecuaciones de verosimilitud que pueden expresarse como sigue

$$\sum_{i=1}^n [y_i - \pi(x_i)] = 0$$

$$y \quad \sum_{i=1}^n x_{ij} [y_i - \pi(x_i)] = 0 \text{ para } j=1, 2, \dots, p$$

También al igual que el caso univariado, la solución de las ecuaciones de verosimilitud requiere métodos especializados los cuales se tienen disponibles en muchos paquetes estadísticos.

2.5 Validación del modelo logístico

Para probar la significancia del coeficiente de una variable en cualquier modelo es necesario plantear la siguiente pregunta, ¿El modelo que incluye la variable en cuestión dice más acerca de la variable dependiente que el modelo que no incluye dicha variable?. Si los resultados obtenidos con la variable en el modelo son mejores, o más exactos en algún sentido, que cuando la variable no está en el modelo, entonces se dice que la variable en cuestión es significativa.

En regresión logística al igual que en el modelo de regresión lineal se comparan los valores observados de la variable dependiente con los valores que predice el modelo con y sin la variable en cuestión. La comparación se basa en el logaritmo del cociente de las funciones de verosimilitud del modelo actual y del modelo saturado³, es decir:

$$D = -2 \ln \left[\frac{(\text{Verosimilitud del modelo actual})}{(\text{Verosimilitud del modelo saturado})} \right] \quad (2.16)$$

La cantidad entre corchetes se conoce como el cociente de verosimilitud. El -2 veces el logaritmo natural es necesario para obtener una cantidad cuya distribución sea conocida y así usarla como prueba de hipótesis. Tal prueba se conoce como la prueba del cociente de verosimilitud. Utilizando el logaritmo de la verosimilitud y operando algebraicamente (2.16) se llega a:

$$D = -2 \sum_{i=1}^n \left[y_i \ln \left(\frac{\hat{\pi}_i}{y_i} \right) + (1 - y_i) \ln \left(\frac{1 - \hat{\pi}_i}{1 - y_i} \right) \right] \quad (2.17)$$

Donde $\hat{\pi}_i = \pi(x_i)$

La ecuación 2.17 es conocida como la devianza y tiene el mismo rol de la suma de cuadrados de los residuales que utiliza la regresión lineal. Para determinar la significancia de una variable independiente se compara el valor de D con y sin la variable independiente en

³ Un modelo saturado es el que contiene tantos parámetros como datos.

la ecuación. El cambio en D debido a la inclusión de la variable independiente en el modelo se obtiene de la siguiente expresión.

$$G = D(\text{para el modelo sin la variable}) - D(\text{Para el modelo con la variable})$$

Utilizando (2.16) se tiene que G se puede calcular de la siguiente manera:

$$G = -2 \ln \left[\frac{\text{Verosimilitud del modelo sin variable}}{\text{Verosimilitud del modelo saturado}} \right] - \left[-2 \ln \left(\frac{\text{Verosimilitud del modelo con variable}}{\text{Verosimilitud del modelo saturado}} \right) \right]$$

Si A= Verosimilitud del modelo sin variable.

B= Verosimilitud del modelo saturado.

C= Verosimilitud del modelo con variable.

entonces

$$G = -2 \ln \left[\frac{A}{B} \right] - \left[-2 \ln \left(\frac{C}{B} \right) \right]$$

$$G = -2(\ln A - \ln B) + 2(\ln C - \ln B)$$

$$G = -2 \ln A + 2 \ln B + 2 \ln C - 2 \ln B$$

$$G = -2 \ln A + 2 \ln C$$

$$G = -2 \ln \left(\frac{A}{C} \right)$$

$$D = -2 \ln \left[\frac{(\text{Verosimilitud sin la variable})}{(\text{Verosimilitud con la variable})} \right]$$

Cuando se tiene el caso simple, es decir con una sola variable independiente es fácil saber que pasa cuando la variable no está en el modelo. La estimación de máxima verosimilitud de b_0 es (n_1/n_0) donde $n_1 = \sum y_i$ y $n_0 = \sum (1 - y_i)$ y el valor que se predice es constante, (n_1/n) . En este caso el valor de G va a estar dado por:

$$G = -2 \ln \left[\frac{\left(\frac{n_1}{n} \right)^{n_1} \left(\frac{n_0}{n} \right)^{n_0}}{\prod_{i=1}^n \hat{\pi}_i^{y_i} (1 - \hat{\pi}_i)^{(1-y_i)}} \right]$$

o bien

$$G = 2 \left\{ \sum_{i=1}^n \left[y_i \ln \left(\hat{\pi}_i \right) + (1 - y_i) \ln \left(1 - \hat{\pi}_i \right) \right] - \left[n_1 \ln(n_1) + n_0 \ln(n_0) - n \ln(n) \right] \right\}$$

Bajo la hipótesis de que b_1 es igual a cero. El estadístico G se distribuye como una variable aleatoria Ji-cuadrada con un grado de libertad.

Otra prueba estadística equivalente que ha sido sugerida por Rao (1973) es la prueba de Wald que se basa en el siguiente estadístico

$$W = \frac{\hat{b}_1}{\hat{SE}(\hat{b}_1)}$$

La prueba se obtiene comparando la estimación de máxima verosimilitud del coeficiente \hat{b}_1 con la estimación de su error estándar. Dicho cociente bajo la hipótesis de que $b_1 = 0$ sigue una distribución normal estándar.

El estadístico G también se puede aplicar cuando el modelo incluye más de una variable independiente, la única diferencia es que los valores ajustados de $\hat{\pi}$, se basan en un vector **B** que contiene p+1 parámetros. Bajo la hipótesis nula de que los p coeficientes del modelo son igual a cero, G también se distribuirá como una Ji-cuadrada con p grados de libertad. El rechazo de la hipótesis nula tiene una interpretación análoga al modelo de regresión lineal múltiple que implica que al menos uno de los coeficientes es diferente de cero.

Para concluir cuales son los coeficientes diferentes de cero, se puede utilizar también la estadística de Wald que se calcula como

$$W_i = \frac{\hat{b}_i}{\hat{SE}(\hat{b}_i)}$$

que también bajo la hipótesis de $b_i = 0$ tiene una distribución normal estándar. El valor del estadístico w_i indicará cuales de las variables son o no significativas al incluirlas en el modelo.

2.6 Bondad de ajuste del modelo

Una manera de determinar que tan bien ajusta un modelo es comparando las estimaciones con los datos observados mediante una tabla de clasificación en la que se distribuyen los casos observados según si la predicción del modelo fue correcta o no.

Cuadro 2.5
Mujeres en edad fértil unidas por uso de métodos anticonceptivos

Valores observados	Predicciones del modelo		Total
	Si usa	No usa	
Si usa	28	12	40
No usa	7	25	32

Fuente: ENADID 1997.

La clasificación de 40 mujeres que son usuarias de métodos anticonceptivos el modelo clasifica 28 correctamente lo que implica un porcentaje de 70 por ciento correctamente clasificado. Por otro lado, de las 32 mujeres que no usan métodos anticonceptivos el modelo clasifica correctamente 25 lo que implica un 70 por ciento de mujeres correctamente clasificadas. En total el porcentaje de clasificación correcta de acuerdo con la tabla es 73.6 por ciento (Cuadro 2.5).

Otra manera de determinar la bondad de ajuste del modelo es analizando que tan probables son los resultados de la muestra dados los parámetros estimados. A la probabilidad de los resultados observados dados los parámetros estimados se le conoce como la verosimilitud. Ya que la verosimilitud es un número menor que uno se acostumbra

usar -2 veces el logaritmo natural de la verosimilitud como una medida de que tan bien el modelo estimado ajusta los datos. Un buen modelo es aquel que tiene una alta verosimilitud lo que se traduce en un valor pequeño de $-2\ln$. Es claro que un modelo que ajusta perfectamente tiene como verosimilitud uno y $-2\ln$ es cero.

Para probar la hipótesis de que la verosimilitud observada no es diferente del valor de la verosimilitud de un modelo que ajusta perfectamente se puede usar el valor $-2\ln$. Bajo la hipótesis de que el modelo ajusta perfectamente $-2\ln$ tiene una distribución Ji-cuadrada con $n-p$ grados de libertad, donde n es el número de casos y p es el número de parámetros estimados.

Otra estadística que se puede usar para probar qué tan bien ajusta el modelo es la que compara las probabilidades observadas y las probabilidades que predice el modelo. La estadística se define como:

$$Z^2 = \sum \frac{\text{residual}^2}{P_i(1-P_i)}$$

Donde el residual es la diferencia entre un valor observado Y_i y el valor predicho P_i . Esta estadística también tiene una distribución Ji-cuadrada si el modelo es correcto, con aproximadamente $n-p$ grados de libertad y se conoce como la Ji cuadrada de Pearson.

Hosmer y Lemeshow en 1980, propusieron formar grupos con base en las probabilidades estimadas. Para esto propusieron dos estrategias de agrupación, la primera de ellas es colapsar la tabla⁴ basándose en percentiles de las probabilidades estimadas y la segunda es colapsarla basándose en valores fijos de la probabilidad estimada.

Con el primer método y usando $g=10$ grupos, el primer grupo contiene los $n'_1 = \frac{n}{10}$ individuos que tienen las probabilidades estimadas más pequeñas, y el último grupo contiene los $n'_{10} = \frac{n}{10}$ individuos que tienen probabilidades estimadas más grandes, con el segundo método y usando $g=10$, los grupos se forman con base en los puntos de corte definidos con los valores $k/10$, con $k=1, 2, 3, \dots, 9$ e integrando los grupos con todos los sujetos

⁴ La tabla se forma con n columnas que corresponden a los n valores de las probabilidades estimadas y con las dos filas correspondientes a los valores de $y=1$ y $y=0$.

cuyas probabilidades estimadas se encuentren entre los puntos de corte adyacentes. Por ejemplo, el primer grupo debe contener todos los sujetos cuyas probabilidades estimadas son menores que 0.1, mientras que el décimo grupo contiene todos los sujetos cuyas probabilidades son menores que 0.9.

Para la fila $y=1$ de la tabla, la estimación del valor esperado se obtiene sumando las probabilidades estimadas sobre todos los sujetos del grupo g respectivo. Para la fila $y=0$, el valor esperado estimado se obtiene sumando sobre todos los sujetos en el grupo g respectivo, el valor 1 menos la probabilidad estimada. La estadística de prueba para ambas estrategias de agrupación está dada por:

$$\hat{c} = \sum_{k=1}^g \frac{\left(o_k - n'_k \bar{\pi}_k \right)^2}{n'_k \bar{\pi}_k \left(1 - \bar{\pi}_k \right)} \quad (2.18)$$

Donde

n'_k es el número de individuos en el k -ésimo grupo,

o_k es la frecuencia de los casos $y = 1$ en el k -ésimo grupo, y

$\bar{\pi}_k$ es el promedio de las probabilidades estimadas en el k -ésimo grupo

Usando un conjunto grande de simulaciones Hosmer y Lemeshow (1980) demostraron que cuando el modelo de regresión logística ajustado es el correcto, la distribución de \hat{c} es aproximadamente Ji-cuadrada con $g-2$ grados de libertad. Los grupos se conocen como los deciles de riesgo.

Se muestran los resultados de aplicar la estrategia de agrupamiento de los deciles de riesgo a las probabilidades calculadas con base en un modelo de regresión logística ajustado a datos de bajo peso (Cuadro 6). En el Cuadro se observa que la frecuencia de bajo peso ($y=1$) para el quinto decil de riesgo es 6, mientras que la frecuencia estimada esperada en este mismo decil de riesgo es de 5.02, la cual es la suma de las probabilidades estimadas

mediante el modelo logístico para los veinte sujetos de ese grupo. La frecuencia observada para el peso normal ($y=0$) es $20-6=14$ y la frecuencia estimada esperada es $20-5.02=14.98$.

Cuadro 2.6
Observaciones y frecuencias estimadas esperadas según decil de riesgo y condición de peso

Peso		Decil de riesgo										Total
		1	2	3	4	5	6	7	8	9	10	
Bajo	Observado	0	1	4	2	6	6	6	10	9	15	59
	Esperado	0.9	1.6	2.3	3.7	5.0	5.6	6.8	8.6	10.5	14.1	59
Normal	Observado	18	19	14	18	14	12	12	9	10	4	130
	Esperado	17.2	18.4	15.8	16.4	15.0	12.4	11.2	10.4	8.5	4.9	130
Total		18	20	18	20	20	18	18	19	19	19	189

Fuente: Hosmer y Lemeshow (1989). Applied Logistic Regression. New York

Con base en los datos del cuadro 6 se calculan los valores de la estadística \hat{c} . Es decir:

para $k=1$

$$\frac{\left(o_1 - n'_1 \bar{\pi}_1\right)^2}{n'_1 \bar{\pi}_1 \left(1 - \bar{\pi}_1\right)} = \frac{(0 - 0.9)^2}{0.9 \left(1 - \frac{0.9}{18}\right)} = 0.9474$$

para $k=2$

$$\frac{\left(o_2 - n'_2 \bar{\pi}_2\right)^2}{n'_2 \bar{\pi}_2 \left(1 - \bar{\pi}_2\right)} = \frac{(1 - 1.6)^2}{1.6 \left(1 - \frac{1.6}{20}\right)} = 0.2446$$

para $k=3$

$$\frac{\left(o_3 - n'_3 \bar{\pi}_3\right)^2}{n'_3 \bar{\pi}_3 \left(1 - \bar{\pi}_3\right)} = \frac{(4 - 2.3)^2}{2.3 \left(1 - \frac{2.3}{18}\right)} = 1.4406$$

Para $k=4, 5, 6, 7, 8, 9$ y 10 los valores son los siguientes $0.9584, 0.2667, 0.0415, 0.1513, 0.4164, 0.4790$ y 0.2228 respectivamente. El valor para la estadística de bondad de ajuste se calcula entonces sustituyendo los valores anteriores en la ecuación 2.18. Es decir,

$$\hat{c} = \sum_{k=1}^g \frac{\left(o_k - n'_k \bar{\pi}_k\right)^2}{n'_k \bar{\pi}_k \left(1 - \bar{\pi}_k\right)} = 5.2$$

El correspondiente valor para un Ji-cuadrada con 8 grados de libertad, es 0.73 , lo que indica que el modelo parece ajustar bastante bien los datos, así como también la comparación de los valores observados y frecuencias esperadas en cada una de las 20 celdas del Cuadro 2.6 muestra que el modelo ajusta bien al interior de cada decil de riesgo.

CAPÍTULO III

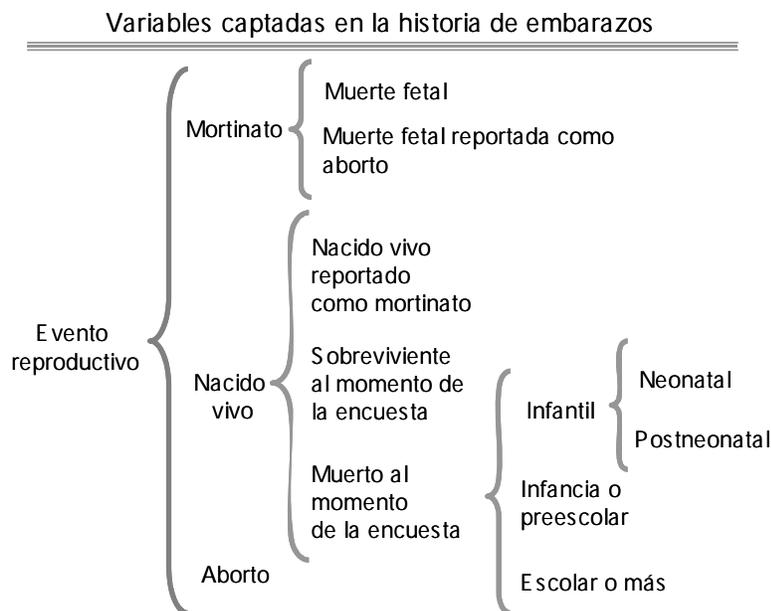
Determinantes de la mortalidad en México

3.1 Tamaño de muestra

Para llevar a cabo el estudio se tomó la información de la historia de embarazos de La Encuesta Nacional sobre Fecundidad y Salud (ENFES) levantada en 1987 por la Secretaría de Salud y la Encuesta de la Dinámica Demográfica (ENADID), levantada en 1997 por el Instituto Nacional de Estadística, Geografía e Informática (INEGI).

La siguiente figura muestra la información que se captó en las respectivas encuestas sobre la historia de embarazos.

Figura 3.1



Fuente: Bobadilla J. Schlaepfer y Alagón J. (1992) Family Formation Patterns and Child Mortality in México.

Es importante aclarar que tanto la ENFES como la ENADID captan información de manera retrospectiva aproximadamente hasta el año de 1960. Sin embargo, algunos autores cuestionan la confiabilidad de la información entre más grande sea el intervalo de tiempo que se considere¹ hacia atrás. Con el fin de contar con tamaños de muestra adecuados, en este trabajo se consideró para ambas encuestas un periodo de 10 años atrás con respecto al año de levantamiento, por lo que en el caso de la ENFES el periodo de análisis es de 1976 a 1986, mientras que en la ENADID es de 1986 a 1996.

Por otra parte, también se observa un traslape de un año en los intervalos, debido a que el levantamiento entre una y otra encuesta es de diez años y a que el análisis en cada una de ellas empieza un año atrás del año de levantamiento. En el Cuadro 3.1 se puede observar la distribución del tamaño de muestra, según las encuestas y los eventos de interés.

¹ El problema reside en que entre más grande es el intervalo, recordar las fechas de los eventos de interés es más difícil.

Cuadro 3.1
Tamaños muestrales según evento de interés y encuesta

Evento	ENFES 1987	ENADID 1997
Nacidos vivos	9,969	74,917
Defunciones menores de un año	486	2,134

Fuente: Cálculos propios en base a la ENFES 1987 y ENADID 1997

Al ponderar los datos del Cuadro 3.1, con la ENFES de 1987, se estiman 20,682,840 nacidos vivos y 1,038,736 defunciones de menores de un año en la década de 1976 a 1986, lo que da como resultado una Tasa de Mortalidad Infantil de 50.2 defunciones de menores de un año por cada mil nacidos vivos, mientras que para la década 1986 a 1996 estimada con la ENADID, fueron 22,056,284 nacidos vivos y 630,653 defunciones de menores de un año lo que implica una Tasa de Mortalidad Infantil de 28.6 defunciones de menores de un año. Se observa que en diez años se logró reducir la tasa en 43 por ciento, es decir, se redujo a casi la mitad la ocurrencia de muertes en menores de un año.

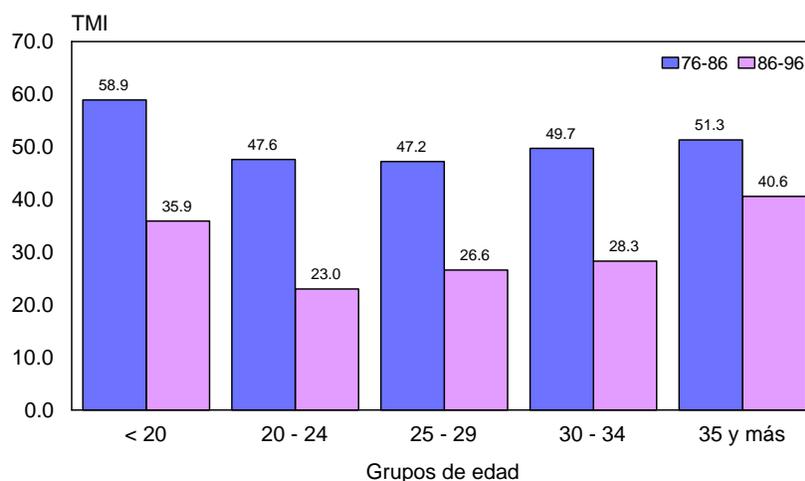
3.2 Análisis comparativo de las variables del patrón familiar

Las estimaciones de las tasas de mortalidad infantil (TMI) por grupos de edad de la madre al momento de tener a su hijo indican que el nivel de la TMI en ambos periodos es casi igual excepto en los grupos de menos de 20 años y de 35 años y más (Gráfica 3.1). En el periodo de 1976 a 1986 se observa que el grupo con más alta TMI era el grupo de mujeres menores de 20 años, mientras que en el periodo de 1986 a 1996 según la ENADID los grupos con más alta tasa fueron el de 35 años y más².

² Con base en una prueba de hipótesis se encontró que las diferencias entre grupos de edad son significativamente diferentes con un nivel de significancia de $\alpha=5\%$. Con la prueba de hipótesis se demuestra que aunque las diferencias sean mínimas, estadísticamente son reales.

Gráfica 3.1

Tasas de mortalidad infantil por grupos de edad de mujeres al nacimiento del hijo, 1976-1986 y 1986-1996

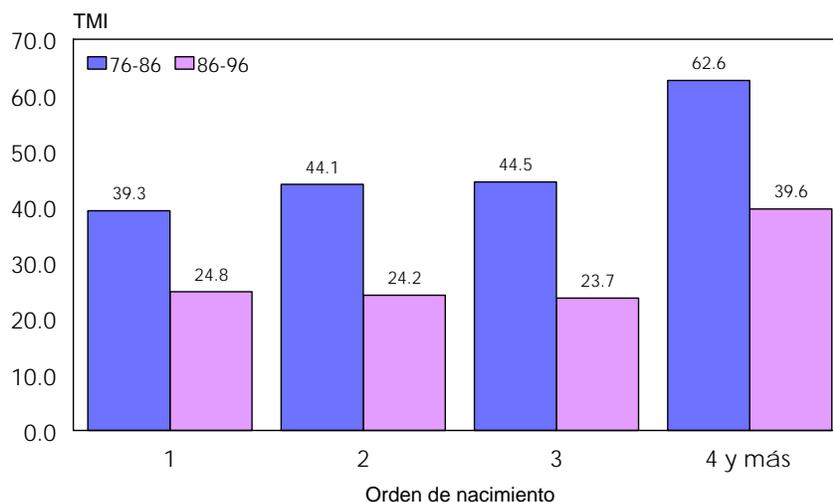


Fuente: cálculos propios con base en la ENFES y ENADID 1997

Las estimaciones de la TMI por orden de nacimiento (Gráfica 3.2) también muestran diferencias estadísticamente significativas³. De acuerdo con los datos se observa en la ENFES una tendencia sostenida a la alza conforme el orden de nacimiento aumenta. En contraste en la ENADID, los datos permanecen casi sin cambio en los ordenes de nacimiento de 1 a 3 y sólo en la última categoría del orden de nacimiento 4 y más se incrementa de manera muy importante.

Gráfica 3.2

Tasas de mortalidad infantil por orden de nacimiento, 1976-1986 y 1986-1996

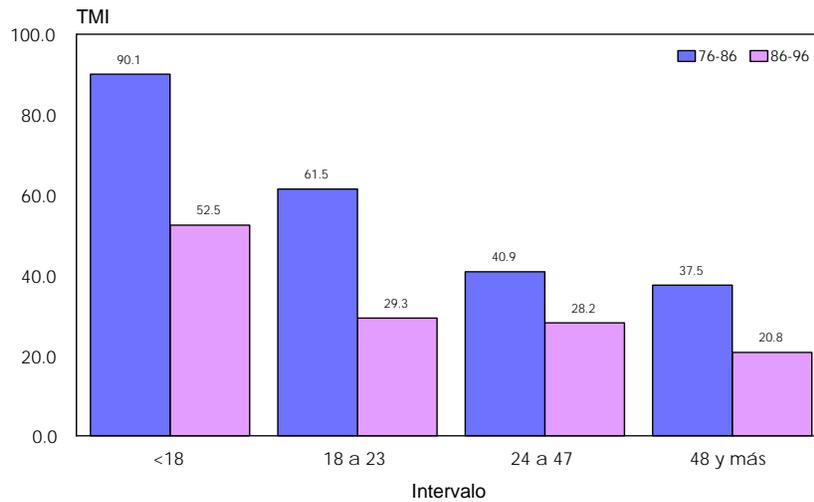


Fuente: cálculos propios con base en la ENFES y ENADID 1997

³ El nivel de significancia de la prueba utilizada fue de $\alpha=5\%$.

Las estimaciones respecto al intervalo intergenésico (Gráfica 3.3) también presentan diferencias significativas. En la ENFES y ENADID se estima que la tasa de mortalidad infantil es más baja entre más grande es el número de meses que hay entre uno y otro nacimiento. En la ENADID las diferencias entre las categorías 18 a 23 y 24 a 27 meses son significativamente diferentes.

Gráfica 3.3
Tasas de mortalidad infantil por intervalo intergenésico,
1976-1986 y 1986-1996

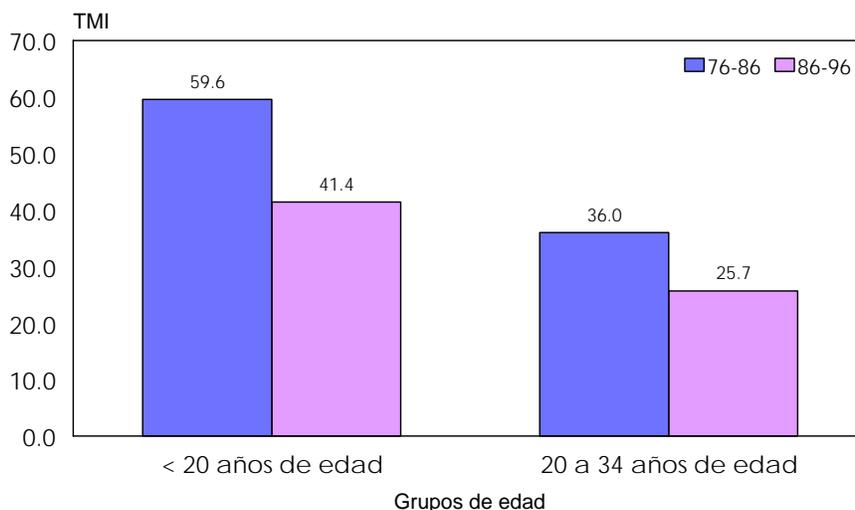


Fuente: cálculos propios con base en la ENFES y ENADID 1997

La tasa de mortalidad infantil en los primeros nacimientos según la edad que tenía la madre (Gráfica 3.4), también presenta marcadas diferencias de acuerdo al grupo de edad de la madre. Siendo más alta en las mujeres cuya edad es menor de 20 años, también al comparar los diferentes periodos se observan cambios muy importantes al bajar de 59.6 a 41.4 en el grupo de menores de 20 años de edad. Por otro lado, en el periodo más reciente las tasas de mortalidad también han bajado aunque en menor medida.

Gráfica 3.4

Tasas de mortalidad infantil en primeros nacimientos por grupos de edad de la madre, 1976-1986 y 1986-1996

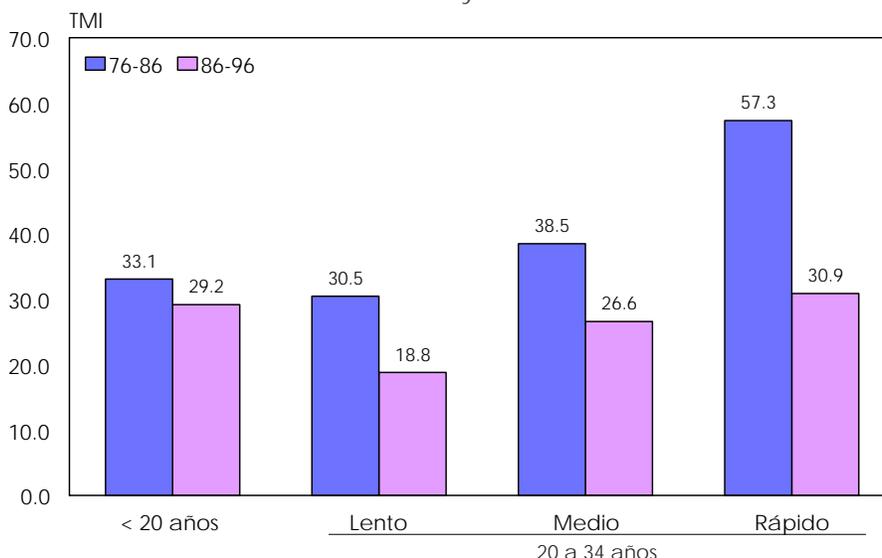


Fuente: cálculos propios con base en la ENFES y ENADID 1997

Otra de las variables importantes del patrón familiar es el ritmo reproductivo, el cual es diferencial cuando se analiza desde el punto de vista de los nacimientos posteriores con intervalo bien espaciado y nacimientos posteriores con intervalo pobremente espaciado. En el primer caso de acuerdo con la ENFES y la ENADID se estima que la tasa de mortalidad infantil se incrementa (Gráfica 3.5). En el segundo caso se observa el mismo comportamiento en ambas encuestas (Gráfica 3.6)

Gráfica 3.5

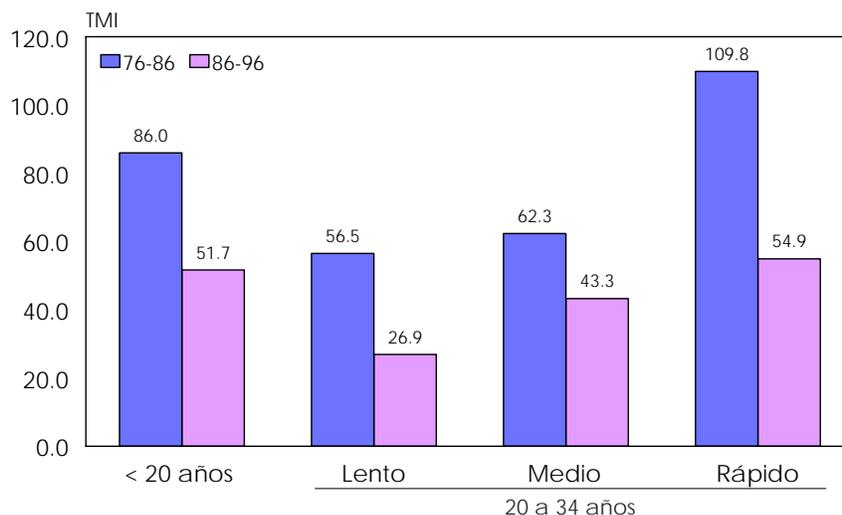
Tasas de mortalidad infantil en nacimientos posteriores con intervalo bien espaciado según ritmo reproductivo, 1976-1986 y 1986-1996



Fuente: cálculos propios con base en la ENFES y ENADID 1997

Gráfica 3.6

Tasas de mortalidad infantil en nacimientos posteriores con intervalo pobremente espaciado según ritmo reproductivo, 1976-1986 y 1986-1996



Fuente: cálculos propios con base en la ENFES y ENADID 1997

3.3 Análisis del patrón familiar desde el enfoque ordinal

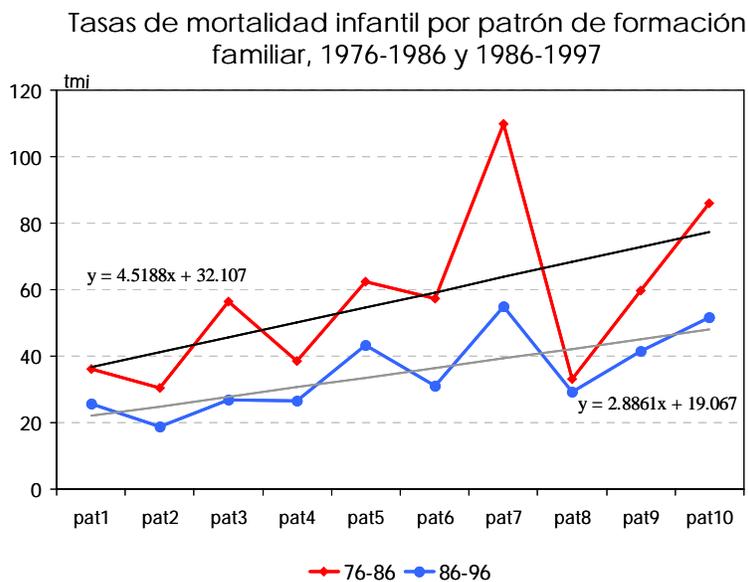
El patrón de formación familiar define diez características de la mujer que pueden considerarse en primer lugar como propicias o no para tener un embarazo, pero además las diferentes características se pueden ordenar de tal manera que el paso de una categoría a otra implique peores condiciones para embarazarse, lo que se traduce a su vez en un mayor riesgo para que el producto de ese embarazo muera en el primer año de vida, por lo que el patrón de formación familiar se puede considerar como una variable ordinal, donde sus categorías implican un mayor riesgo de mortalidad infantil.

El análisis de la mortalidad infantil visto desde el enfoque ordinal del patrón de formación familiar, indica que a mayor valor de esta variable, en general la tasa de mortalidad infantil también es mayor en ambos periodos del análisis (Gráfica 3.7), lo que se puede interpretar como que, entre más adversas sean las condiciones de la mujer para embarazarse, el riesgo de morir en el primer año de vida del niño o niña producto de ese embarazo, es mayor.

Se observa que para la década comprendida entre el periodo de 1976-1986, las estimaciones con base en la encuesta de la ENFES de 1987, muestran que cuando el nivel

del patrón de formación familiar es el más bajo, la tasa de mortalidad infantil es de 36 defunciones por cada mil nacidos vivos, mientras que cuando el nivel es el más alto, la tasa de mortalidad es de 86 defunciones, lo que muestra un incremento de un poco más del doble. Para el caso de la década de 1986-1996, de acuerdo con los datos de la ENADID de 1997, se estima para el nivel más bajo del patrón de formación familiar una tasa de mortalidad infantil de 25.7 defunciones por cada mil nacidos vivos y para el valor más alto del patrón una tasa de mortalidad infantil de 51.7, que representa un incremento prácticamente del doble.

Gráfica 3.7



Fuente: cálculos propios con base en la ENFES y ENADID 1997

También se puede observar que al ajustar una recta para ambos periodos las pendientes han cambiado (4.5 en 1976-1986 y 2.9 en 1986-1996). Esto implica que el impacto del patrón de formación familiar en la mortalidad infantil se ha reducido por cada categoría del patrón. Es decir, en la década 1976-1986 cada vez que la mujer incrementaba sus condiciones adversas del embarazo, la mortalidad infantil se incrementaba en 4.5 defunciones, mientras que para la década de 1986-1996 el incremento es de 2.9, lo que indica una reducción del 64 por ciento.

Otro dato interesante que se obtiene del ajuste de un modelo lineal, es la mortalidad infantil asociada con aquel patrón de formación familiar ideal en el que no existan riesgos de

embarazo (nivel cero). En la década 1976-1986 las mujeres que se embarazaban en condiciones ideales, la mortalidad infantil asociada era de 32.1 defunciones por cada mil nacidos vivos. Para la década siguiente, las estimaciones indican una tasa de mortalidad infantil de 19.1, es decir una reducción del 40.5 por ciento, con respecto a la década anterior.

También es importante resaltar que aún logrando reducir los riesgos asociados con el patrón de formación familiar, la mortalidad infantil en México (20.9 defunciones por cada mil nacidos vivos)⁴ presenta todavía niveles de países como Albania, Bahamas, República de Moldova y Mauritania, cuyos niveles de mortalidad infantil en 1996 eran de 20.4, 18.4, 20.4 y 19.0 respectivamente. Pero también es importante destacar que aún en esas condiciones México todavía se encuentra muy lejos de los niveles de los países como Estados Unidos, Canadá, Francia, Alemania y Suecia cuyos niveles de mortalidad infantil actuales son de 6.5, 4.75, 4.26, 4.16 y 2.77, respectivamente. Con respecto a Latinoamérica nuestros niveles de mortalidad infantil todavía están muy lejos de países como Costa Rica, Chile, Puerto Rico y Cuba, cuyos niveles de mortalidad infantil hoy en día son de 9.95, 8.80, 8.24 y 6.33, respectivamente.

Lo anterior implica que debe haber otras condiciones aparte del patrón de formación familiar, que también tienen un efecto en la mortalidad de los niños menores de un año. El efecto de estas variables o condiciones de la mujer explicaría los niveles de mortalidad todavía altos en nuestro país, aún considerando que la mujer se embaraza en condiciones ideales según el patrón de formación familiar.

3.4 Variables biosocioeconómicas

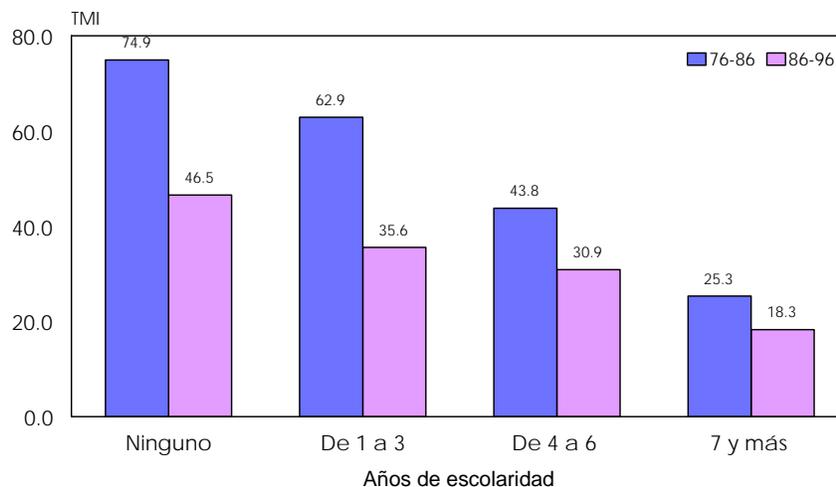
Diversas condiciones socioeconómicas de los mexicanos y las mexicanas afectan la mortalidad infantil. Entre otras se encuentran la escolaridad, el género, la condición laboral, el lugar de residencia y las condiciones de la vivienda.

Los datos de la ENFES y de la ENADID indican que a mayor escolaridad de la mujer y del cónyuge la mortalidad infantil es más baja, por ejemplo la tasa de mortalidad infantil en el

⁴ Index mundi, Tasas de mortalidad infantil –Comparación entre países en 2005.

periodo 1976 a 1986 se estimó en 74.9 por mil nacidos vivos para las mujeres con ningún grado de escolaridad, mientras que para aquellas mujeres cuyo nivel de escolaridad es de 7 años y más se estima en 25.3 por mil nacidos vivos lo que implica una reducción de 66 por ciento en la tasa (Gráfica 3.8). Diez años después se observa que el patrón no ha cambiado aunque la brecha ha disminuido, de acuerdo con la ENADID se estima que en el periodo 1986 a 1996 la reducción es de 61 por ciento, ya que se observa una tasa de mortalidad infantil de 46.5 en hijos de mujeres con ningún año de escolaridad a 18.3 en hijos de mujeres que cuentan con 7 años o más de escolaridad.

Gráfica 3.8
Tasa de mortalidad infantil por años de escolaridad de la madre, 1976-1986 y 1986-1996

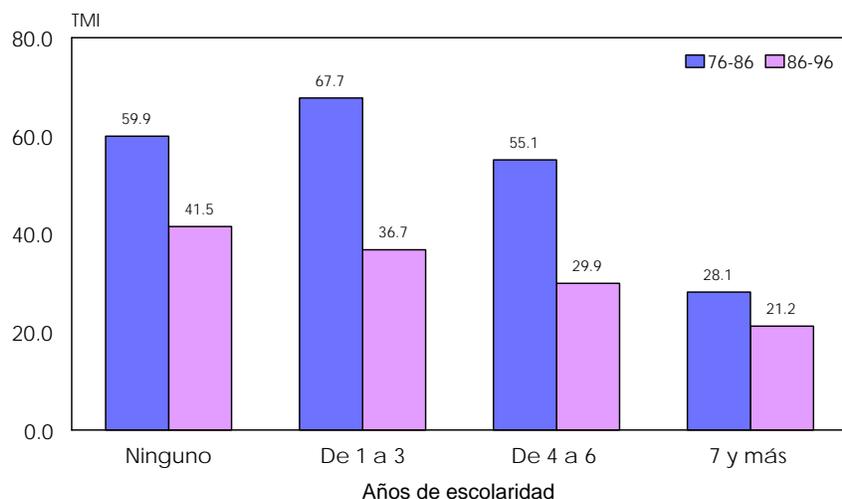


Fuente: cálculos propios con base en la ENFES y ENADID 1997

El análisis de la mortalidad infantil según la escolaridad del cónyuge también muestra el mismo patrón anteriormente observado. Sin embargo, las brechas entre el nivel de escolaridad de ninguno y siete años y más son menos amplias. En el periodo 1976 a 1986 se estimó una diferencia, según la ENFES, en 53 por ciento, ya que se estima en 59.9 defunciones de menores de un año por cada mil nacidos vivos cuando el cónyuge no cuenta con años de escolaridad y si cuenta con 7 años o más la tasa se estima en 28.1, mientras que en el periodo 1986 a 1996 de acuerdo con la ENADID la brecha se reduce a 49 por ciento aunque la tasa haya disminuido a 41.5 cuando el cónyuge no cuenta con años de escolaridad y a 21.2 cuando cuenta con 7 años o más de escolaridad (Gráfica 3.9).

Gráfica 3.9

Tasa de mortalidad infantil por años de escolaridad del cónyuge, 1976-1986 y 1986-1996

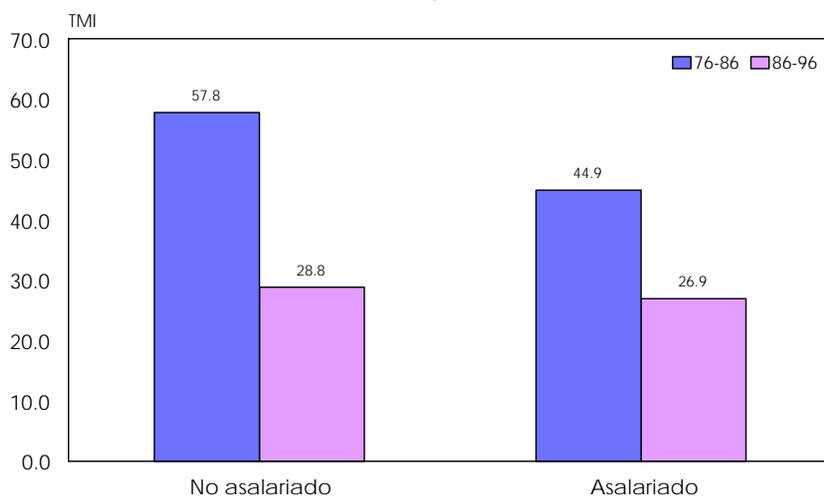


Fuente: cálculos propios con base en la ENFES y ENADID 1997

Otra condición del cónyuge que afecta la mortalidad infantil es su condición laboral, ya que la mortalidad infantil es más alta en aquellos cuya situación laboral es no asalariados. En el periodo de 1976 a 1986 la ENFES arroja cifras de 57.8 para los no asalariados y de 44.9 para los asalariados. Para el periodo 1986 a 1996 la brecha casi se ha reducido, ya que según la ENADID la tasa de mortalidad infantil se estima en 28.8 cuando el cónyuge es no asalariado y de 26.9 si es asalariado, la diferencia entre estas dos condiciones laborales es de sólo 1.9 defunciones (Gráfica 3.10).

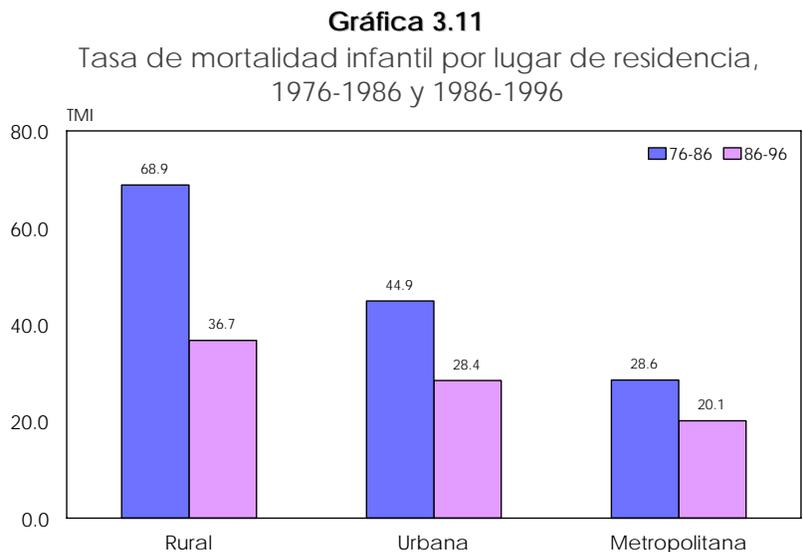
Gráfica 3.10

Tasa de mortalidad infantil por condición laboral del cónyuge, 1976-1986 y 1986-1996



Fuente: cálculos propios con base en la ENFES y ENADID 1997

Por otro lado, las estimaciones indican que las condiciones que rodean a los hogares también afecta a la mortalidad infantil en México. De acuerdo con los datos de la ENFES y la ENADID el lugar de residencia está fuertemente asociado con la mortalidad infantil. Los datos de ambas encuestas indican que entre más desarrollado es el lugar de residencia más baja es la mortalidad infantil. En el periodo 1976 a 1986 la ENFES estima una tasa de mortalidad entre el área rural de 68.9, en áreas urbanas de 44.9 y en áreas metropolitanas de 28.6 defunciones, siendo la diferencia más amplia entre las áreas rural y urbana con 24 defunciones y entre urbana y metropolitana de 16.8, mientras que en el periodo 1986 a 1996 según la ENADID se estimó una tasa de 37.7 en áreas rurales, 28.4 en urbanas y 20.1 en áreas metropolitanas, reduciéndose la brecha entre las áreas de residencia a 9.3 y 8.3 defunciones respectivamente (Gráfica 3.11).

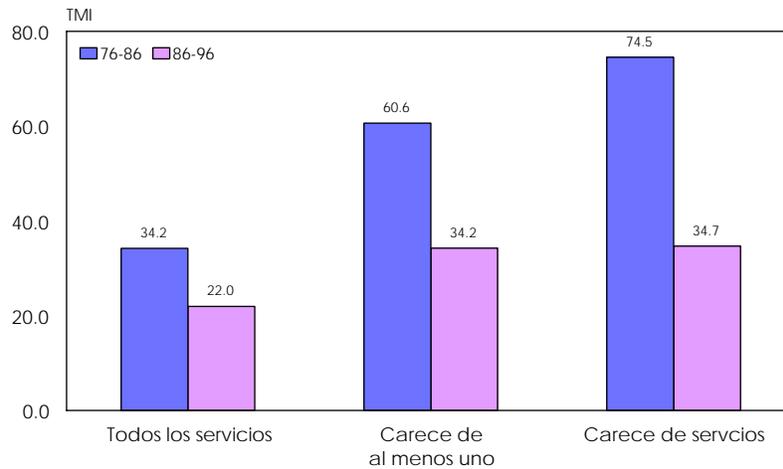


Fuente: cálculos propios con base en la ENFES y ENADID 1997

Otra característica de desarrollo económico que afecta a la mortalidad infantil es la tenencia de servicios en la vivienda. Al parecer entre más servicios tiene la vivienda la mortalidad infantil se reduce. De acuerdo con la ENFES, se estima una tasa de mortalidad infantil de 34.2 cuando la vivienda cuenta con todos los servicios, aumenta a casi el doble 60.6 si carece de al menos un servicio y a 74.5 si carece de todos los servicios. De acuerdo a la ENADID la tasa de mortalidad infantil disminuyó a 22.0 defunciones si la vivienda cuenta con todos los servicios, a 34.2 cuando carece de al menos un servicio y a 34.7 si la vivienda carece de todos los servicios 34.7 (Gráfica 3.12)

Gráfica 3.12

Tasa de mortalidad infantil por tenencia de servicios* en la vivienda, 1976-1986 y 1986-1996



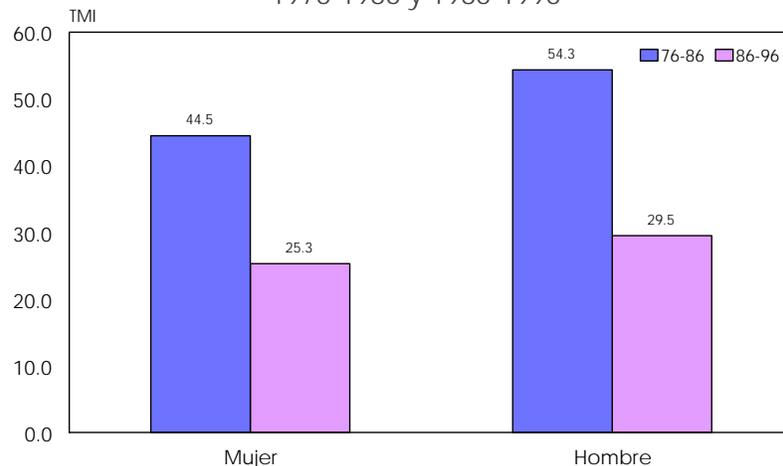
Fuente: cálculos propios con base en la ENFES y ENADID 1997

*Servicios: Agua entubada, electricidad, drenaje y piso diferente de tierra

Finalmente se observa que la mortalidad infantil también es diferencial de acuerdo con el sexo del recién nacido. De acuerdo con la ENFES, la tasa de mortalidad infantil de mujeres es de 44.5 defunciones mientras que para los hombres es de 54.3 defunciones, la diferencia entre ambos sexos es de 9.8 defunciones. En el periodo de 1986 a 1996 disminuye la tasa de mortalidad en ambos casos, así como la brecha entre los mismos al pasar a 4.2 defunciones de diferencia, ya que en mujeres se estimó una tasa de 25.3 defunciones de menores de un año y de 29.5 defunciones en hombres (Gráfica 3.13).

Gráfica 3.13

Tasa de mortalidad infantil por sexo del niño, 1976-1986 y 1986-1996



Fuente: cálculos propios con base en la ENFES y ENADID 1997

3.5 Análisis multivariado de la mortalidad infantil

El análisis descriptivo de la sección anterior muestra estadísticamente que cada una de las variables está muy relacionada con la ocurrencia de la muerte de niños menores de un año. Sin embargo, con este tipo de análisis sólo es posible saber si dos o quizá hasta tres variables influyen simultáneamente en la ocurrencia del evento, pero como se ha visto, dicho evento se encuentra determinado por un gran número de factores, por lo que, para determinar el efecto de cada variable en la probabilidad de muerte, fue necesario hacer el análisis con base en la teoría del modelo de regresión logística.

Es importante señalar que en la estimación del modelo de regresión logística se incluyen las variables biosocioeconómicas junto con el patrón de formación familiar. El patrón de formación familiar por sí sólo explica el comportamiento reproductivo de la mujer y con las demás variables se incluye una parte del entorno social, económico y biológico que incide en la salud de los niños y que se captaron en las dos encuestas que se utilizan en este trabajo.

Si la mujer o el cónyuge cuentan con algún nivel de escolaridad les permite tener acceso a información que les ayude a planear mejor el número y tiempo para tener al o los hijos que deseen, esto con el fin de dar una atención de calidad al o los hijos, si el cónyuge es asalariado probablemente se encuentra afiliado a alguna institución de salud que le permite atender a tiempo cualquier enfermedad o mantener un chequeo médico en forma constante para evitar la enfermedad. El lugar en el que reside es importante porque determina el acceso a los servicios de salud como clínicas u hospitales privados o públicos, así como el abastecimiento de medicamentos en farmacias. Si la vivienda en la que reside el niño no cuenta con los principales servicios como son agua entubada, electricidad, drenaje o piso diferente de tierra, el niño corre el riesgo de vivir en condiciones precarias que pueden afectar su salud. Por último el sexo del infante, estadísticamente se ha comprobado que nacen más niños pero son más vulnerables biológicamente en el primer año de vida.

Debido a que el evento muerte sólo se puede determinar por la ocurrencia o no ocurrencia del mismo, se definió la variable dependiente dicotómica asignando los valores 0 (muerte) y 1 (sobrevive). En el Cuadro 3.1 se presentan todas las variables independientes que se

propusieron para explicar la sobrevivencia de los niños y en el Cuadro 3.2 se encuentran las variables que mejor la explican, según el procedimiento de selección del mejor modelo.

Cuadro 3.1
Variables independientes para el modelo de regresión logística

Variable	Tipo	Códigos
Patrón de formación familiar	Ordinal	1-10
Escolaridad de la madre	Ordinal	0. Ninguno 2. De 4 a 6 años 1. De 1 a 3 años 3. 7 y más años
Sexo del niño	Categórica	0. Mujer 1. Hombre
Características de la vivienda (indiviv0)	Categórica	0. No carece 1. Carece de todos los servicios
Características de la vivienda (indiviv1)	Categórica	0. No carece 1. Carece de al menos un servicio
Escolaridad del cónyuge	Ordinal	0. Ninguno 2. De 4 a 6 años 1. De 1 a 3 años 3. 7 y más años
Condición laboral del cónyuge	Nominal	1. Asalariado 0. Si no
Zona rural	Nominal	1. Rural 0. Si no
Zona urbana	Nominal	1. Urbana 0. Si no

Fuente: ENFES, 1987 y ENADID 1997.

Según datos de la ENFES de 1987, para la década de 1976-1987 con base en el modelo de regresión logística ajustado, se encontró que sólo dos variables explican de manera significativa la ocurrencia de la sobrevivencia infantil (Cuadro 3.2). En primer lugar obsérvese que el coeficiente del patrón de formación familiar es negativo, lo que indica que entre más adversas son las condiciones para tener un embarazo, la probabilidad de sobrevivir del niño o niña es más baja. Otra variable que resultó significativa es la escolaridad de la madre. Las estimaciones muestran que por el contrario del patrón de formación familiar, entre mayor es la escolaridad de la madre la probabilidad de sobrevivir del niño o niña también es mayor.

Con este resultado del modelo se puede decir que la escolaridad y la condición laboral del cónyuge, el lugar de residencia, las características de la vivienda y el sexo del niño no eran un factor determinante para que sobreviviera el niño o niña, pero que además el análisis univariado no detecta. Hay que recordar por otro lado que hasta 1974 se modifica la

política poblacionista de 1947⁵, donde se busca regular racionalmente y estabilizar el crecimiento de la población, para alcanzar la meta se lleva tiempo en implementar los programas de planeación familiar para que haya un impacto en la población, por lo que el modelo refleja como el comportamiento reproductivo (patrón de formación familiar) de esa década de análisis (1976-1986) era el factor a modificar para lograr una mayor sobrevivencia en los niños o niñas, y se demuestra que la educación en ese momento era fundamental ya que los niños de mujeres con mayor escolaridad tenían más probabilidad de sobrevivir en esa década.

Cuadro 3.2
Variables significativas que explican la ocurrencia o no del evento y su cociente de momios

1976-1986			1986-1996		
Variable	B	Cociente de momios	Variable	B	Cociente de momios
Patrón de formación familiar	-0.1086	0.8971	Patrón de formación familiar	-0.0851	0.9184
Escolaridad de la madre	0.2644	1.3027	Escolaridad de la madre	0.1812	1.1986
Constante	3.0466		Hombres	-0.1663	0.8468
			Carencia de al menos un servicio en la vivienda	-0.1987	0.8198
			Constante	3.6844	

Fuente: Cálculos propios con base en la ENFES de 1987 y la ENADID de 1997.

En términos del análisis del cociente de momios los datos indican que cada vez que la mujer incrementa su nivel de escolaridad, la propensión de sobrevivencia se incrementa 1.3 veces, de tal manera que cuando una mujer pasa del nivel de escolaridad más bajo (sin años de escolaridad) al más alto (siete años o más de escolaridad), la propensión a que sobreviva se incrementa casi cuatro veces. En contra parte los datos muestran que, cada vez que la mujer pasa de un nivel de patrón de formación familiar a otro, por ejemplo, primeros nacimientos de mujeres entre 20 a 34 años a nacimientos posteriores de mujeres entre 20 a 34 años con ritmo reproductivo lento e intervalo largo, la propensión a sobrevivir se reduce en 0.90 veces.

⁵ El objetivo de la Ley General de Población de 1947 era privilegiar a familias grandes, fomentar el aumento de nacimientos y disminución de la mortalidad, protección biológica y legal de la infancia mejoras a la alimentación, higiene de casas y centros de trabajo.

Para la década de 1986-1996, las estimaciones obtenidas con base en la ENADID 1997 muestran cambios en la mortalidad infantil. En primer lugar se encontró que dos nuevas variables resultaron altamente significativas para explicar la muerte o sobrevivencia del niño o la niña (Cuadro 3.2). Dichas variables son el sexo del producto del embarazo y la carencia de al menos un servicio en la vivienda.

En cuanto al patrón de formación familiar y escolaridad de la madre no se encontraron cambios significativos, ya que según los datos de la ENADID por cada nivel de escolaridad que incrementa la mujer, la propensión de sobrevivir se incrementa en 1.2, mientras que la reducción en la propensión de sobrevivencia aumenta cuando la mujer cambia de un nivel de patrón de formación familiar a otro se reduce la probabilidad de sobrevivencia en aproximadamente 0.92.

Por otro lado, se encontró que la probabilidad de sobrevivencia cuando el producto del embarazo es hombre es 16 por ciento más baja que la mujer, mientras que cuando existe al menos la carencia de un servicio básico, la probabilidad de sobrevivencia es más baja en aproximadamente 20 por ciento que cuando la vivienda tiene todos los servicios.

En la década de 1986 a 1996, siguen sin ser factores determinantes para que sobrevivan los niños o niñas, la escolaridad y posición laboral del cónyuge y el lugar de residencia.

3.6 Validación del modelo

Existen tres elementos que se pueden utilizar para evaluar la bondad de ajuste del modelo de regresión logístico multivariado. La estadística de Wald, la estadística de Hosmer y Lemeshow y la tabla de clasificación.

Para el caso de los datos correspondientes a la década de 1976-1986, la estadística de Hosmer y Lemeshow se estimó en 13.37 y con base en una Ji-cuadrada con 8 grados de libertad se encontró que la significancia del modelo fue de 0.0996, lo que permite concluir con un nivel de significancia de 10 por ciento, que al menos una de las siete variables que se propusieron para explicar la sobrevivencia infantil es diferente de cero.

Se puede observar que del total de siete variables propuestas sólo el patrón de formación familiar y la escolaridad de la madre resultaron altamente significativas según la estadística de Wald. También es importante señalar la presencia altamente significativa de una constante en el modelo (Cuadro 3.3).

Cuadro 3.3
Estadística de Wald para las variables significativas, 1976-1986

Variable	Coefficiente	Desviación estándar	Wald	Grados de libertad	Significancia
Patrón de formación familiar	-0.1086	0.0185	34.5359	1	0.00000
Escolaridad de la madre	0.26444	0.0523	25.5481	1	0.00000
Constante	3.0466	0.1426	456.649	1	0.00000

Fuente: Cálculos propios con base en la ENFES 1987.

Al analizar el porcentaje de clasificación debido al modelo de regresión logística para el periodo 1976-1986 (Cuadro 3.4). Se observa que de un total de 7,720 casos, 403 que no sobrevivieron, el modelo los clasificó como sobrevivientes, lo que implica un porcentaje de mala clasificación de 5.2 por ciento. En contra parte se observa que 7,317 casos que si sobrevivieron, el modelo los clasificó como sobrevivientes, por lo que el porcentaje de clasificación acertada es de 94.8 por ciento. Estos tres elementos permiten concluir que en general el modelo de regresión logística ajustado explica de manera aceptable la ocurrencia de la sobrevivencia del niño o niña en México en la década de 1976-1986.

Cuadro 3.4
Tabla de clasificación de observaciones, 1976-1986
Predicciones

Observada	Predicciones	
	No sobrevive	Si sobrevive
No sobrevive	0	403
Si sobrevive	0	7,317
Total	0	7,720

Por otro lado, para el periodo 1986-1996, los datos con base en la ENADID de 1997 muestran que la estadística de Hosmer y Lemeshow es de 15.55 por lo que de acuerdo con una chi-cuadrada con ocho grados de libertad se calculó que el modelo es significativo con un

nivel de cinco por ciento, lo que implica que al menos una de las siete variables que se propusieron en el Cuadro 3.1, explica la probabilidad de la sobrevivencia infantil.

Cuadro 3.5
Estadística de Wald para las variables significativas, 1986-1996

Variable	Coefficiente	Desviación estandar	Wald	Grados de libertad	Significancia
Patrón de formación familiar	-0.0851	0.0084	103.4546	1	0.00000
Escolaridad de la madre	0.1812	0.0255	50.2987	1	0.00000
Sexo	-0.1663	0.0482	11.8881	1	0.00006
Carencia de al menos un servicio	-0.1987	0.0516	14.7975	1	0.00001
Constante	3.6844	0.0840	1925.727	1	0.00000

Para validar el modelo ajustado en el periodo de 1986 a 1996 se presenta información en el Cuadro 3.5. De acuerdo con estos datos para la década de 1986 a 1996, las variables que resultaron altamente significativas de acuerdo con la estadística de Wald fueron, el patrón de formación familiar, las escolaridad de la madre, el sexo y la carencia de al menos un servicio básico. También se puede notar la significancia importante de una constante en el modelo. El nivel de significancia de las variables y la constante es de al menos siete por cien mil.

Cuadro 3.6
Tabla de clasificación, 1986-1996
Predicciones

Observada	Predicciones	
	No sobrevive	Si sobrevive
No sobrevive	0	1,799
Si sobrevive	0	57,124
Total	0	58,923

Con respecto al porcentaje de clasificación del modelo de regresión logística en esta década, en el Cuadro 3.6 se puede observar que de un total de 58,923 casos, el modelo

clasifica erróneamente a un 3.1 por ciento. Por otro lado el modelo clasifica correctamente el 96.9 por ciento. Nuevamente con base en estos tres elementos, se puede concluir que el modelo de regresión logística explica de manera aceptable la ocurrencia de la sobrevivencia del niño o niña en México en la década correspondiente a los años de 1986-1996.

Resumen y conclusiones

Con respecto al patrón de formación familiar

En cuanto a la edad de la mujer al momento de tener a su hijo se estima que existen diferenciales importantes de un grupo a otro y que no se observan cambios en la tendencia, en ambos periodos. La mortalidad infantil empieza alta para las mujeres jóvenes, después desciende para las mujeres de 20 a 34 años y sube nuevamente para las mujeres añosas.

Los datos de la mortalidad infantil según el orden de nacimiento, indican que la mortalidad infantil se incrementa sustancialmente en ambos periodos cuando el orden de nacimiento es de cuatro y más. Aunque en el periodo de 1986-1996 se observa una estabilización para los ordenes uno, dos y tres.

La tasa de mortalidad infantil por intervalo intergenésico también muestra diferencias importantes, a medida que el intervalo es más grande. En ambos

periodos la reducción es de más de la mitad al pasar de un intervalo de menos de 18 meses a uno de 48 meses y más.

Con respecto a la mortalidad infantil según los primeros nacimientos y edad de la madre, el patrón es el mismo en ambos periodos. Alta tasa de mortalidad infantil para las mujeres jóvenes y baja para las mujeres con edad entre 20 y 34 años.

En cuanto a nacimientos posteriores de mujeres con intervalo bien espaciado se observa en ambos periodos un incremento de la tasa de mortalidad conforme el ritmo se incrementa. Sin embargo, de un periodo a otro la tasa de mortalidad infantil, ha bajado sustancialmente, además en el periodo de 1986-1996 las diferencias entre las tasas según las categorías que componen al ritmo reproductivo también bajaron.

La mortalidad infantil en nacimientos posteriores de mujeres con intervalos pobremente espaciados, presenta el mismo patrón de comportamiento que para las mujeres con intervalo bien espaciado. Al comparar la mortalidad entre este tipo de mujeres se observan grandes brechas ya que el nivel de mortalidad para las mujeres con intervalo pobremente espaciado es casi el doble que el de las mujeres con intervalos bien espaciados.

Con respecto al patrón de formación familiar se concluye que el impacto al pasar de condiciones más desfavorables para el embarazo se ha reducido, en la década de 1976-1986 el incremento era de 4.5 defunciones, mientras que para la década de 1986-1996 el incremento de pasar de una condición de adversidad a otra fue de 2.9 es decir, una reducción de 64 por ciento.

El análisis del patrón de formación familiar considerando que representa condiciones más o menos desfavorables para tener un embarazo, indica que el efecto de las condiciones más desfavorables en la mortalidad infantil se ha reducido a la mitad.

Con respecto a las variables biosocioeconómicas

El análisis de las variables biosocioeconómicas muestra que la escolaridad, el género, la condición laboral, el lugar de residencia y la tenencia de servicios en la vivienda tienen un efecto importante en la mortalidad infantil.

Los datos de escolaridad indican para ambos periodos que cuando la mujer tiene un nivel de escolaridad alto, la mortalidad infantil se reduce. Para el caso de la escolaridad del cónyuge se tiene el mismo patrón pero sólo para aquellos cónyuges cuyo nivel es de uno a siete años y más.

Por condición laboral del cónyuge, se observa una reducción muy importante entre los no asalariados y los asalariados para el periodo 1976-1986, mientras que para la década de 1986-1996 no es tan importante.

El lugar de residencia de la mujer también es diferencial para la mortalidad infantil. En ambos periodos se observa que al pasar de un área rural a una metropolitana la mortalidad infantil se reduce de manera importante. Sin embargo, las brechas de un lugar de residencia a otro se han reducido.

Con respecto a la tenencia de servicios en la vivienda, el patrón de mortalidad infantil en ambos periodos es el mismo. Baja mortalidad en viviendas con todos los servicios y alta en viviendas que carecen de ellos. Finalmente se puede concluir que el sexo sigue siendo un factor que produce diferencias en la mortalidad. Siendo en el hombre más alta que en la mujer.

Con respecto al análisis multivariado

Con base en el análisis multivariado de la mortalidad infantil se concluye que algunos de los efectos de las variables del patrón de formación familiar y de las biosocioeconómicas desaparecen al considerar de manera simultánea todo este

tipo de variables. Por ejemplo, de las nueve variables propuestas para explicar la mortalidad infantil se encontró en la década de 1976-1986 que sólo dos de ellas explican la mortalidad infantil, mientras que para la década de 1986-1996 son cuatro las variables. Esta diferencia en el número de variables, probablemente se deba a los tamaños muestrales manejados en las respectivas encuestas. Ya que el tamaño de muestra en la ENADID es casi ocho veces más grande que la encuesta de la ENFES.

Para la década de 1976-1986 los datos muestran que el patrón de formación familiar tiene un efecto en la mortalidad infantil negativo, lo que permite concluir que entre más adversas son las condiciones para tener un embarazo, la probabilidad de sobrevivir del infante es más baja. La otra variable que resultó importante en este mismo periodo es la escolaridad de la madre, los datos indican que tiene un efecto contrario al patrón de formación familiar, ya que en este caso entre mayor es la escolaridad de la madre, la probabilidad de sobrevivir del niño o niña también aumentan.

Para la década de 1986-1996 de acuerdo con los datos de la ENADID se concluye que tanto el patrón de formación familiar como la escolaridad de la madre siguen explicando la mortalidad infantil, pero sus efectos casi no han cambiado.

Por otro lado, en el periodo de 1986 a 1996 se observa que cuando el producto del embarazo es hombre, su riesgo de sobrevivir es 16 por ciento más bajo que cuando es mujer. También se observa que cuando existen carencias de servicios básicos en la vivienda, la probabilidad de sobrevivir del niño o niña, es más baja en aproximadamente 20 por ciento, al compararse con viviendas que tienen todos los servicios.

Los datos de validación de ambos modelos muestran en general significancias muy altas. Es importante reconocer que para el modelo aplicado a los datos de la década de 1976-1986, la significancia del modelo fue de diez por ciento, que es considerada por algunos investigadores como una significancia baja. Pero esto

se debe a que la muestra de la ENFES es pequeña para analizar este fenómeno, y porque además el cálculo del tamaño de muestra no se hizo pensando en llevar a cabo este tipo de análisis, de ahí que la muestra no haya sido lo suficientemente grande, para analizar la mortalidad infantil en México.

En el caso de la ENADID aunque la muestra no se calculó para hacer análisis de la mortalidad infantil, si se calculó para que tuviera representatividad a nivel de entidad federativa, lo que implica que los análisis a nivel nacional tienen una muestra muy grande, que permiten el análisis de fenómenos como la mortalidad infantil y otros.

En resumen de acuerdo con los datos de validación se puede concluir que los modelos de regresión logística, explican de manera aceptable la probabilidad de sobrevivir de los niños o niñas menores de un año en México, en las décadas correspondientes a 1976-1986 y 1986-1996 y que por lo tanto las conclusiones que se realicen con base en los modelos, pueden considerarse confiables para la toma de decisiones con respecto al fenómeno de la mortalidad infantil en México.

Con respecto al nivel global de la mortalidad infantil

La mortalidad infantil en México bajó de manera muy importante entre 1981 y 1991. Datos de ENFES 1987 indican que la mortalidad infantil era de 50.22 defunciones por cada mil nacidos vivos para 1981, diez años después se estima según la ENADID de 1997 en 28.59 defunciones lo que implica una reducción de 43.1 por ciento.

Sin embargo, las cifras de la mortalidad infantil en México, todavía se encuentran muy lejos de los niveles alcanzados en países como Estados Unidos, Francia, Alemania y Suecia, donde la mortalidad infantil se estima en 6.50, 4.75, 4.26, 4.16, y 2.77 defunciones por cada mil nacidos vivos respectivamente, por lo que trabajos de investigación como esta tesis, pueden contribuir en la definición de

estrategias que reduzcan la mortalidad de los niños y niñas menores de un año en México.

Finalmente la contribución de este trabajo de investigación en estudios de la mortalidad infantil, es la incorporación del patrón de formación familiar como una variable ordinal, ya que hasta hoy ningún trabajo relacionado con este tipo de investigaciones lo había contemplado de esta manera. Este enfoque resuelve el problema de colinealidad de las diferentes categorías del patrón, al intentar incorporarlas como variables independientes entre sí en la explicación de la mortalidad infantil. Por otra parte, se puede concluir que no es del todo adecuado considerar a esta variable como una medición de los patrones de formación familiar, sino más bien para medir condiciones adversas en las que la mujer pueda embarazarse.

Anexo

Cuadro 1
Marco determinantes próximos y variables intermedias de la fecundidad y de la sobrevivencia en la infancia

Determinantes Próximos	Variables intermedias de sobrevivencia de la infancia	Variables de fecundidad
<i>Factores de exposición a la concepción</i>		
1. Unión sexual	Edad materna, paridad	Unión Marital
2. Frecuencia del coito	Intervalo intergenésico	Fecundidad
<i>Factor de lactancia</i>		
3. Lactancia materna	Ingesta dietética, intervalo intergenésico	Amenorrea de la lactancia
<i>Factor de riesgo ecológico</i>		
4. Deficiencia dietética	Consumo de calorías, proteínas y micronutrientes.	Fecundidad (hambruna), fecundidad o esterilidad (enfermedad venérea, tuberculosis, entre otras)
5. Contaminación ambiental	Contaminación del aire, agua, alimentos, dedos, piel, tierra, objetos, insectos, vectores o bien incidencia, prevalencia de infecciones respiratorias, diarrea, parasitosis intestinal, tétanos, infecciones de la piel, paludismo, entre otras	Abortos espontáneos (paludismo y otro)
6. Accidentes	Lesiones accidentales	Fecundidad y/o esterilidad, abortos espontáneos (lesiones al nacimiento)

Cuadro 1
Marco determinantes próximos y variables intermedias de la fecundidad
y de la sobrevivencia en la infancia

Determinantes Próximos	Variables intermedias de sobrevivencia de la infancia	Variables de fecundidad
<i>Factor de riesgo ecológico</i>		
7. Deficiencia dietética	Consumo de calorías, proteínas y micronutrientes.	Fecundidad (hambruna), fecundidad o esterilidad (enfermedad venérea, tuberculosis, entre otras)
8. Contaminación ambiental	Contaminación del aire, agua, alimentos, dedos, piel, tierra, objetos, insectos, vectores o bien incidencia, prevalencia de infecciones respiratorias, diarrea, parasitosis intestinal, tétanos, infecciones de la piel, paludismo, entre otras	Abortos espontáneos (paludismo y otro)
9. Accidentes	Lesiones accidentales	Fecundidad y/o esterilidad, abortos espontáneos (lesiones al nacimiento)
<i>Factores de intervención directa</i>		
10. Medidas preventivas personales	Inmunizaciones, atención prenatal, atención del parto	Contracepción, esterilización
11. Medidas curativas	Tratamientos usados	Fecundidad
12. Lesiones intencionalmente inflingidas	Infanticidio, circuncisión femenina.	Aborto inducido.

Fuente: Mosley /Chen, Bongaart

Bibliografía

Abbagnano Nicola (1998). *Diccionario de filosofía*, México.

Aguirre A. (1982). *Atención materno-infantil y su relación con la mortalidad materna e infantil : un análisis regional*, INEGI, México.

(1997). *Cambios en la mortalidad infantil*, UNAM, Instituto de Investigaciones Sociales, México.

Bobadilla J. , Schlaepfer L. y Alagón J. (1990). *Family Formation Patters and Child Mortality in Mexico*, The Population Council, New York.

Bobadilla J. L. y Langer A. (1990). *Mortalidad infantil en México: un fenómeno en transición*, UNAM, Instituto de Investigaciones Sociales, México.

Bongaarts P. John (1982) "*The proximate determinants of natural marital fertility*". New York: The Population Council, p 1-44: il.

Bronfman M. Gómez de León J. (1984). *La Mortalidad en México, niveles tendencias y determinantes*, El Colegio de México, México.

Cervantes A. (1989). *En busca de la explicación, variables socioeconómicas en el estudio de la mortalidad en la niñez*, UNAM, Centro Regional de Investigaciones Multidisciplinarias, México.

CONAPO, (2002). *Proyecciones de población 2000-2030*, México.

Dallas E. J. (2000). *Métodos multivariados aplicados al análisis de datos*, Thomson Editores, México.

Department for Economic and Social Information and Policy Analysis Population Division
United Nations (1994). *The Health Rationale for Family Planning: Timing of Births
and Child Survival*, New York.

Fernández S. B. Comp. (1994). *La salud de los niños : cifras nacionales*, Nueva York:
UNICEF.

Ferrater M. J. (1994). *Diccionario de Filosofía*, México.

Freund R. J. and Wilson W.J. (1993). *Statistical Methods*, Academic Press, Inc, San Diego,
California, USA.

Gómez de León J., Franco F. y Potter J. (1993). *Factores determinantes de la mortalidad
infantil en México: un modelo de efectos aleatorios*, SSA, Subsecretaría de
Coordinación y Desarrollo, México.

Hernández H. (1998). *Algunos determinantes de la mortalidad infantil aún sin mayor
bienestar la mortalidad infantil disminuye*, UNAM, Instituto de Investigaciones
Sociales, México.

Hosmer D. y Lemeshow S. (1989). *Applied Logistic Regression*, Massachussets, USA.

Index mundi, (2005). *Tasas de mortalidad infantil –Comparación entre países*.

Instituto Nacional de Estadística Geografía e Informática (1999). *Encuesta Nacional de
la Dinámica Demográfica 1987*, México.

Jiménez R., Vera G. y Ruiz G. (1986). *Factores socio-demográficos asociados a la
mortalidad infantil en México un enfoque empírico*, El Colegio de México,
UNAM-SOMEDE, México.

John E. F., Irwin M. y Marylees M., (2000). *Estadística Matemática con Aplicaciones*.

- Juárez F. (1992). *Intervención de las instituciones en la reducción de la fecundidad y la mortalidad infantil*, El Colegio de México, México.
- Levy P.S and Lemeshow S. (1991). *Sampling of Populations: Methods and Applications*, Interscience Publication John Wiley And Sons, Inc. New York.
- Mier M. y Terán (1986). *Consecuencias en la salud de las tendencias en la fecundidad y formación de familias en México*, El Colegio de México, UNAM-SOMEDE, México
- Minujin, A. (1984). *Factores sociodemográficos asociados a la mortalidad Infantil*, El Colegio de México, México.
- Mojarro O., Aznar R. (1986). *Influencia de los factores biológicos y de la estructura social en la mortalidad infantil, 1965-1974* , IMSS, México.
- Mosley, W.H. and Lincon C. Chen (1984). *An Analytical Framework for the Study of the Child Survival in Developing Countries*, *Population and Development Review*, Supplement 10, New York, USA.
- Naciones Unidas (1994). *The Health Rationale for Family Planning Timing of Births and Child Survival*, New York.
- Ornelas R. Coord. (1989). *Investigación multidisciplinaria de la mortalidad y morbilidad en niños menores de cinco años: memorias del primer seminario de demografía formal*. UNAM, Centro Regional de Investigaciones Multidisciplinarias, México.
- Partida B. V. (2005). *La mortalidad en los primeros años de vida*, México ante los Desafíos de Desarrollo del Milenio. SEGOB-CONAPO, México.

Preston S. H., Patrick H. and Michel G. (2002). *Demography Measuring and Modeling Population Processes*, Massachusetts, USA.

Rao, C.R. (1973). *Linear Statistical Inference and Its Application*. Second edition. Wiley, New York

Rasmussen S. (1991). *An introduction to statistics with data analysis*, Brooks/ Cole Publishing Company Pacific Grove, California, USA.

Reynolds H. T. (1977). "The Analysis of Cross-Classifications" .

Rohde J, Allman J. " *Infant mortality in relation to fertility control practice in developing countries*". International Population Conference 1981; 2: 97-112.

Secretaría de Salud, Subsecretaría de Servicios de Salud y Dirección General de Planificación Familiar. (1989). *Encuesta Nacional sobre Fecundidad y Salud 1987*, México

United Nations (1997). *Demographic Yearbook, 49th Issue General Tables*, New York, USA.

Wackerly D., Mendenhall W y Scheaffer R. L. (2002). *Estadística Matemática con Aplicaciones*, Thomson Editores, México.

Cuadros y Gráficas

Cuadro 1.1 Ritmo reproductivo según edad de la madre y número de hijos.....	10
Cuadro 2.1 México. Condición de sobrevivencia al primer año de vida por sexo.....	22
Cuadro 2.2 Expresiones del modelo logístico simple según variable dependiente e independiente.....	24
Cuadro 2.3 México. Condición de sobrevivencia al primer año de vida por lugar de residencia.....	26
Cuadro 2.4 Variables dummy de lugar de residencia de la mujer.....	27
Cuadro 2.5 Mujeres en edad fértil unidas por uso de métodos anticonceptivos.....	38
Cuadro 2.6 Observaciones y frecuencias estimadas esperadas según decil de riesgo y condición de peso.....	41
Cuadro 3.1 Tamaños muestrales según evento de interés y encuesta.....	45
Gráfica 3.1 Tasas de mortalidad infantil por grupos de edad de mujeres al nacimiento del hijo, 1976-1986 y 1986-1996.....	46
Gráfica 3.2 Tasas de mortalidad infantil por orden de nacimiento, 1976-1986 y 1986-1996.....	46
Gráfica 3.3 Tasas de mortalidad infantil por intervalo intergenésico, 1976-1986 y 1986-1996.....	47
Gráfica 3.4 Tasas de mortalidad infantil en primeros nacimientos por grupos de edad de la madre, 1976-1986 y 1986-1996.....	48
Gráfica 3.5 Tasas de mortalidad infantil en nacimientos posteriores con intervalo bien espaciado según ritmo reproductivo, 1976-1986 y 1986-1996.....	48
Gráfica 3.6 Tasas de mortalidad infantil en nacimientos posteriores con intervalo pobremente espaciado según ritmo reproductivo, 1976-1986 y 1986-1996.....	49
Gráfica 3.7 Tasas de mortalidad infantil por patrón de formación familiar, 1976-1986 y 1986-1997.....	50
Gráfica 3.8 Tasa de mortalidad infantil por años de escolaridad de la madre, 1976-1986 y 1986-1996.....	52

Gráfica 3.9 Tasa de mortalidad infantil por años de escolaridad del cónyuge, 1976-1986 y 1986-1996.....	53
Gráfica 3.10 Tasa de mortalidad infantil por condición laboral del cónyuge, 1976-1986 y 1986-1996.....	53
Gráfica 3.11 Tasa de mortalidad infantil por lugar de residencia, 1976-1986 y 1986-1996.....	54
Gráfica 3.12 Tasa de mortalidad infantil por tenencia de servicios en la vivienda, 1976-1986 y 1986-1996.....	55
Gráfica 3.13 Tasa de mortalidad infantil por sexo del niño, 1976-1986 y 1986-1996.....	55
Cuadro 3.1 Variables independientes para el modelo de regresión logística.....	57
Cuadro 3.2 Variables significativas que explican la ocurrencia o no del evento y su cociente de momios.....	58
Cuadro 3.3 Estadística de Wald para las variables significativas, 1976-1986.....	60
Cuadro 3.4 Tabla Clasificación de observaciones, 1976-1986.....	60
Cuadro 3.5 Estadística de Wald para las variables significativas, 1986-1996.....	61
Cuadro 3.6 Tabla de clasificación de observaciones, 1986-1996.....	61