



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

**MODELO LINEAL GENERALIZADO PARA
ÍNDICES DE POBREZA EXTREMA**

TESIS

QUE PARA OBTENER EL TÍTULO DE:

A C T U A R I A

P R E S E N T A

ELISANA PEREA ROMERO

Tutor

DR. JOSÉ RAMIRO SÁNCHEZ AGUILAR

2013





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.

1. Datos de la alumna

Perea
Romero
Elisana
58468506
Universidad Nacional Autónoma de México
Facultad de Ciencias
402014755

2. Datos del Tutor

Doctor
José Ramiro
Sánchez
Aguilar

3.- Sinodal 1

Doctora
María del Pilar
Alonso
Reyes

4. Sinodal 2

Doctora
Laura Elena
Gloria
Hernández

5. Sinodal 3

Doctora
María Cristina
Gutiérrez
Delgado

6. Sinodal 4

Doctora
Claudia Gisela
Vázquez
Cruz

Título

Modelo lineal generalizado para índices de pobreza extrema
77 Páginas
2013

Dedicatorias

A mi familia:

Por ser los mejores amigos; porque en todo momento me alientan a seguir luchando para alcanzar mis sueños y proporcionándome su apoyo incondicional en cada empresa que abordo.

A mis amigos:

Por ser mis compañeros incansables que siempre han apoyado mis proyectos más ambiciosos.

A mis profesores:

Por ser la guía que me llevó a concluir esta parte de mi carrera profesional.

Agradecimientos

A mi familia:

Por el apoyo al alentar y fomentar mi crecimiento personal y profesional.

A mi tutor Dr. Ramiro Sánchez y cotutor Act. Fabián González:

Por la oportunidad de culminar esta fase de mi carrera y por el entusiasmo, dedicación y apoyo que proporcionan a todos los proyectos de Actuario que la Universidad forma.

A mis entrañables amigas:

Angie Godínez, Maricela Solano e Irais Bravo, por los ánimos y apoyo que me dieron durante este proyecto.

A la C.P. Laura Liyen y la Lic. Rosario Salinas:

Quienes me mostraron que la Universidad es y será siempre un trampolín a las estrellas.

Índice general

Índice general.....	I
Índice de cuadros	III
Índice de figuras	V
Introducción.....	1
Capítulo 1 Panorama de la pobreza en México: factores, causas y consecuencias	3
Introducción.....	3
1.1 Diagnóstico situacional de la pobreza en México	3
1.1.1 Definición unidimensional de la pobreza	8
1.2 Los factores no económicos que determinan la pobreza.....	10
1.3 Análisis cuantitativo de la pobreza	12
1.3.1 Para las entidades federativas con mayor proporción de personas en situación de pobreza en México	14
1.3.2 Por zona de residencia (rural - urbano).....	19
1.4 Mediciones de la pobreza en México	22
1.4.1 Índice de pobreza de acuerdo a ingreso anterior a 2004	22
1.4.2 Metodología multidimensional	23
1.5 Políticas públicas para mitigar la pobreza en el país	23
1.5.1 Plan Nacional de Desarrollo (2007 – 2012)	23
1.5.2 Plan Visión 2030	25
1.5.3 Programa de apoyo alimentario en zonas de atención prioritaria	25
1.5.4 Programa de abasto social de leche.....	26
1.5.5 Programa de desarrollo Oportunidades.....	26
1.5.6 Comisión Nacional de Derechos Humanos	27
Capítulo 2 Modelo lineal generalizado para la estimación de la pobreza extrema.....	29
Introducción.....	29
2.1 Uso de un modelo lineal generalizado para estimar la proporción de la pobreza extrema.....	29
2.2 Notación y nomenclatura	30
2.3 Planteamiento teórico: Modelos Lineales Generalizados	32
2.3.1 Ventajas de los Modelos Lineales Generalizados sobre los Modelos Lineales ...	32
2.3.2 Componentes de los Modelos Lineales Generalizados	33
2.3.3 Modelos para variables respuesta con dos categorías.	35
2.3.4 Verificación de la adecuación del modelo.....	38

Capítulo 3 Modelo de estimación de probabilidad de pobreza extrema	41
Introducción	41
3.1 Propósito y alcance	41
3.2 Diseño metodológico para determinar índices de pobreza extrema.....	42
3.3 Diseño, planeación y ejecución.....	44
3.3.1 Ejecución del modelo	53
3.3.2 Validación de la aplicación del modelo	62
3.3.3 Ajuste del modelo.....	68
3.4 Resultados.....	68
Conclusiones	71
Referencias Bibliográficas	75
Glosario	77

Índice de cuadros

Cuadro 1.1 Enfoque multidimensional de la pobreza	5
Cuadro 1.2 Medición de la Pobreza, Estados Unidos Mexicanos, 2010	13
Cuadro 1.3 Medición de la Pobreza, Chiapas, 2010.....	14
Cuadro 1.4 Medición de la Pobreza, Guerrero, 2010.....	15
Cuadro 1.5 Medición de la Pobreza, Oaxaca, 2010.....	16
Cuadro 1.6 Medición de la Pobreza, Tlaxcala, 2010.....	17
Cuadro 1.7 Medición de la Pobreza, Zacatecas, 2010.....	18
Cuadro 1.8 Municipios con mayores porcentajes de población en situación de pobreza	20
Cuadro 1.9 Municipios con mayor cantidad de gente en situación de pobreza.....	21
Cuadro 2.1 Notación para las distribuciones de probabilidad	31
Cuadro 2.2 Algunas funciones liga y sus inversas.....	34
Cuadro 3.1 Definición de valores por indicador de pobreza.....	45
Cuadro 3.2 Fuentes de información por indicador	46
Cuadro 3.3 Origen dela información disponible	47
Cuadro 3.4 Información disponible por año de levantamiento.....	48
Cuadro 3.5 Resultados de la aplicación del MLG con 7 variables - 2004	54
Cuadro 3.6 Resultados del análisis de bondad de ajuste al MLG con 7 variables - 2004	55
Cuadro 3.8 Resultados de la aplicación del MLG con 8 variables - 2006	57
Cuadro 3.9 Resultados de la aplicación del MLG eliminando variables - 2006.....	57
Cuadro 3.10 Aplicación del modelo lineal generalizado Logit - 2008	59
Cuadro 3.11 Aplicación del modelo lineal generalizado Probit - 2008	59
Cuadro 3.12 Aplicación del modelo lineal generalizado Logit - 2010	60
Cuadro 3.13 Aplicación del modelo lineal generalizado Probit - 2010	60
Cuadro 3.14 Aplicación del modelo lineal generalizado Logit - 2012	61
Cuadro 3.15 Aplicación del modelo lineal generalizado Probit - 2012	61
Cuadro 3.16 Indicadores de validez de modelos aplicados – 2004.....	62
Cuadro 3.17 Indicadores de validez de modelos aplicados - 2006	63
Cuadro 3.18 Indicadores de validez de modelo Logit - 2008	64
Cuadro 3.19 Indicadores de validez de modelo Probit - 2008	64
Cuadro3.20 Indicadores de validez del modelo logit-2010.....	65

Cuadro3.21 Indicadores de validez del modelo Probit-2010	66
Cuadro3.22 Indicadores de validez del modelo Logit-2012	66
Cuadro3.23 Indicadores de validez del modelo Probit-2012	67
Cuadro 3.24 Coeficientes finales de la modelación	68
Cuadro 3.25 Comparación de la proporción de la población en situación de pobreza extrema de acuerdo al modelo lineal generalizado respecto a estimación CONEVAL	69

Índice de figuras

Figura 1.1 México, Evolución de la pobreza según tres versiones de la Línea de Bienestar, 1968 – 2002 (Porcentaje de personas pobres)	6
Figura 1.2 Porcentaje de personas en pobreza por la dimensión de ingreso 1992 - 2010.....	7
Figura 1.3 Correlación entre los factores de la pobreza	12
Figura 1.4 Medición de pobreza 2010 por municipio. Porcentaje de la población en situación de pobreza por municipio, México, 2010.....	20
Figura 1.5 Porcentaje de la población en situación de pobreza extrema por municipio, México, 2010	21

Introducción

El objetivo de este estudio es construir un modelo lineal generalizado a partir de la evaluación que hace el CONEVAL e identificar las entidades federativas con mayor índice de pobreza extrema, con el propósito de incorporar el enfoque multidimensional que establece la Ley General de Desarrollo Social para la estimación del índice de pobreza extrema.

En 2004 el Congreso de la Unión decretó la Ley de Desarrollo Social con el objetivo de garantizar los derechos sociales consagrados en la Constitución, así como establecer y delimitar funciones y obligaciones del Gobierno para encausar los esfuerzos para lograr tal fin. Adicionalmente, la Ley de Desarrollo Social establece que el Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL) deberá evaluar a la pobreza en intervalos de dos años por entidad federativa y cada cinco años a nivel municipal; sin embargo el Ejecutivo Federal debe dirigir esfuerzos y recursos anualmente para disminuir la pobreza extrema en las zonas de atención prioritaria donde lo amerite.

En el artículo 36 de la Ley General de Desarrollo Social, de aplicación nacional se define a la pobreza y su método de medición, el cual debe utilizar cuando menos los siguientes indicadores: ingreso corriente per cápita, rezago educativo, acceso a servicios de salud, acceso a la seguridad social, calidad de la vivienda, acceso a servicios en la vivienda, acceso a alimentación y hacinamiento. Así, el CONEVAL presentó una metodología, definiendo a una persona en situación de pobreza cuando presenta al menos una carencia social y un ingreso menor al costo de las necesidades básicas. La población en pobreza extrema presenta tres o más carencias sociales y su ingreso es menor al valor de la canasta alimentaria.

En la última evaluación realizada por el CONEVAL (2010) se reportó que los dos rubros que tuvieron un detrimento y, por tanto, aumentaron el indicador de pobreza en México, fueron: el poder adquisitivo del ingreso y la carencia a la alimentación; en consecuencia el 28.7% de la población es vulnerable por carencia social; 5.8% es vulnerable por ingreso; 35.8% son pobres moderados y 10.4% de la población padece pobreza extrema. En este panorama los estados con mayor acumulación de pobreza son Chiapas, Oaxaca, Puebla, Guerrero, Tlaxcala y Zacatecas.

La metodología propuesta por CONEVAL considera variables dicotómicas o categóricas, restringiendo la evaluación de la correlación entre los indicadores, por lo que resulta conveniente utilizar Modelos Lineales Generalizados que permitan: evaluar la dependencia entre variables, facilitar la segmentación de variables o la inclusión de nuevas variables

relevantes para la medición de la pobreza (nivel educativo, hacinamiento promedio por vivienda, rango de ingreso por persona, espacio de la vivienda por citar algunos ejemplos) y construir de proyecciones a corto, mediano y largo plazo, para establecer un plan de desarrollo sostenido.

La tesis se presenta grosso modo en tres capítulos:

El objetivo del capítulo 1 es delimitar el panorama actual de la población mexicana que vive en pobreza, los factores que la afectan y las consecuencias que afectan la realidad nacional, destacando las regiones donde la atención de la pobreza, por su gravedad es prioritaria y son objeto de atención del gobierno. También se presentan las entidades con mayores índices de pobreza.

Por su parte, el objetivo del capítulo 2 es estudiar los Modelos Lineales Generalizados; presentar y argumentar el uso de un modelo lineal generalizado para estimar la pobreza extrema considerando, cuando menos, el uso de las variables establecidas en la Ley General de Desarrollo Social. También se realiza un análisis de la información de distintas bases de datos para realizar las estimaciones a futuro.

Finalmente en el capítulo 3 se construirá un modelo con la información correspondiente a la población mexicana para inferir las proporciones de ésta que cuentan con alguna combinación de factores que inducirían a la pobreza extrema.

En las conclusiones se presentará una síntesis de los resultados, las dificultades presentadas a lo largo de la ejecución y los hallazgos obtenidos en éste trabajo. Las propuestas y áreas de oportunidad de la información o bien, de otras formas de abordar y aportar al problema de medición son presentadas en este apartado.

Capítulo 1

Panorama de la pobreza en México: factores, causas y consecuencias

Introducción

La pobreza ha sido un problema presente en todas las sociedades ricas y pobres a lo largo de la historia, como consecuencia de la incapacidad de los gobiernos para otorgar oportunidades justas a los habitantes de sus países, por la presencia de guerras, catástrofes naturales o por las crisis globales que afectan a las naciones involucradas. En general, la pobreza se convierte en un fenómeno a abordar con distintos enfoques los cuales proporcionan una visión integral de su origen.

La pobreza puede existir siempre que no cause dolor en la sociedad que la padece. En México, la pobreza es constante y se agudiza con las crisis económicas recurrentes, sean nacionales o internacionales.

En este capítulo se presenta la información respecto a la pobreza, el sustento legal y metodológico con el que se define para su aplicación en la población mexicana, sus causas, afectaciones y evolución para la gente que vive bajo esta situación y las implicaciones para el resto de la sociedad.

1.1 Diagnóstico situacional de la pobreza en México

A través de los años se desarrollaron distintas concepciones de pobreza, consecuencia de la definición en sí. En las acepciones iniciales se denotaba a una persona como pobre cuando carecía de medios necesarios para subsistir o bien el capital primordialmente monetario para la adquisición de los bienes necesarios. Sin embargo, desde la Revolución Industrial, algunos autores han destacado en sus discursos que la pobreza no debe conceptualizarse en un sentido unidimensional.

Adam Smith, en su discurso Investigación de la naturaleza y causas de la riqueza de las naciones¹ afirmó que adicional a las necesidades primarias del hombre (alimento, vivienda) existían necesidades de acuerdo al contexto cultural y social, y si un hombre careciese de

¹ Smith, Adam, Investigación de la naturaleza y causas de la riqueza de las naciones (1794), Edición digital de la Universidad de Sevilla.

estos bienes o servicios, se consideraría a sí mismo pobre, quedando en desventaja respecto a los demás, sin ser necesariamente pobre extremo. A esta discrepancia asociada a factores distintos a las necesidades en sí se le considera pobreza relativa y cambiante de acuerdo a la cultura y momento en el tiempo.

Los autores clásicos (Marx, Malthus y Senior entre los más destacados) criticaron fuertemente las iniciativas de protección a los desvalidos y pobres, pues argumentaban que la ayuda proporcionada resultaba en un lastre para la sociedad. Dadas las críticas se delimitó el concepto de pobreza en términos multidimensionales con el objeto de acotar el alcance para así, otorgar la ayuda sin fomentar el incremento de la pobreza.

De los estudios y las disertaciones efectuadas por siglos se tienen tres enfoques: en derechos, de bienestar y de acuerdo al contexto territorial.

El enfoque de derechos se asocia a la existencia de garantías fundamentales, inalienables, insustituibles e interdependientes; por lo que la pobreza consiste en la negación de los derechos. Algunos estudios con éste enfoque consideran a la pobreza como una afectación directa a todos los derechos humanos.

El enfoque de bienestar tiene por objetivo identificar las condiciones que limitan la libertad de las personas para tener un desarrollo pleno. En este mismo enfoque se considera la disponibilidad de recursos económicos para obtener condiciones de desarrollo mediante la adquisición de bienes y servicios.

Ambos enfoques, derechos y bienestar, son atributos personales o de la unidad doméstica. Sin embargo, la pobreza ha mostrado que cuenta con elementos comunitarios asociados al territorio y, donde la mitigación o disminución pasa a ser responsabilidad del Estado.

Atendiendo a la necesidad de contar con una definición alineada al contexto mexicano, el Congreso de la Unión, con el sustento de los artículos 3 y 4 de la Constitución Política de los Estados Unidos Mexicanos², decretó la Ley General de Desarrollo Social³ que los indicadores mínimos a utilizar para la medición multidimensional de la pobreza establecidos por el CONEVAL en el artículo 36 son de aplicación obligatoria para las entidades y dependencias públicas que participen en la ejecución de los programas de desarrollo social, y deberá utilizar la información que genere el Instituto Nacional de Estadística y Geografía,

² Constitución Política de los Estados Unidos Mexicanos, Ciudad de México, Diario Oficial de la Federación, 5 de febrero de 1917, Última reforma publicada 26 de febrero, 2013.

³ Ley General de Desarrollo Social, Ciudad de México, Diario Oficial de la Federación, 16 de junio 2012.

independientemente de otros datos que se estime conveniente, al menos sobre los siguientes indicadores:

- I. Ingreso corriente per cápita;
- II. Rezago educativo promedio en el hogar;
- III. Acceso a los servicios de salud;
- IV. Acceso a la seguridad social;
- V. Calidad y espacios de la vivienda;
- VI. Acceso a los servicios básicos en la vivienda;
- VII. Acceso a la alimentación, y
- VIII. Grado de cohesión social.

Es así como los enfoques para la pobreza empatan con la descripción de los indicadores, estableciendo un marco de medida completo para el caso mexicano.

Cuadro 1.1
Enfoque multidimensional de la pobreza

Enfoque		
Bienestar	Derechos	Contexto territorial
Ingreso corriente per cápita	Rezago educativo promedio en el hogar Acceso a los servicios de salud Acceso a la seguridad social Calidad y espacios de la vivienda Acceso a los servicios básicos en la vivienda Acceso a la alimentación	Grado de cohesión social

Fuente: Elaboración propia.

Al establecerse la Ley⁴, también se designó al Consejo Nacional para la Evaluación de la Política de Desarrollo Social (CONEVAL) como el encargado de medir la pobreza. En el estudio Lineamientos y criterios generales para la definición, identificación y medición de la pobreza⁵ se define el término pobreza multidimensional como aquella cuyos ingresos son insuficientes para adquirir los bienes y los servicios que requiere para satisfacer sus necesidades y presente carencia en al menos uno de los siguientes seis indicadores: rezago educativo, acceso a los servicios de salud, acceso a la seguridad social, calidad y espacios de la vivienda, servicios básicos en la vivienda y acceso a la alimentación.

⁴ Ídem.

⁵ Consejo Nacional de Evaluación de la Política de Desarrollo social, *Lineamientos y criterios generales para la definición, identificación y medición de la pobreza*, Ciudad de México, Diario Oficial de la Federación, 2010.

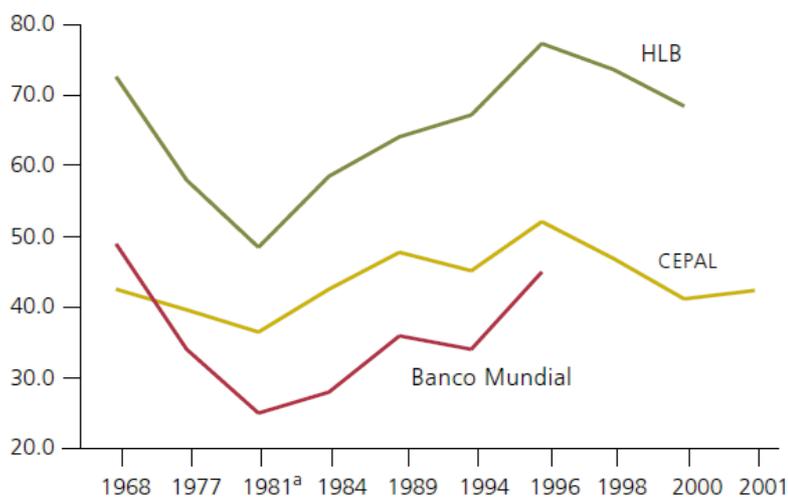
A su vez, en el mismo estudio se define la población en pobreza extrema:

“Una persona se encuentra en situación de pobreza multidimensional extrema cuando no tenga garantizado el ejercicio de al menos tres de los derechos constitutivos para el desarrollo social y sus ingresos sean inferiores a la línea de bienestar mínimo.”

La medición de la pobreza en México, previo a la publicación de la Ley General de Desarrollo Social, fue variada debido a las diferencias en los métodos para medirla y los agudos contrastes entre los umbrales utilizados. De acuerdo a lo descrito por Araceli Damian y Julio Boltvinik⁶, la mayoría de las series de evolución de la pobreza disponibles están basadas en el método de línea de pobreza.

Una fuente de diferencias importantes con las que se enfrenta el investigador es la información estadística existente pues ocasionalmente no es compatible y de manera adicional, los levantamientos de distintos años tienen objetivos distintos.

Figura 1.1
México, Evolución de la pobreza según tres versiones de la Línea de Bienestar, 1968 – 2002
(Porcentaje de personas pobres)



Fuente: Damian, Araceli y Boltvinik, Julio, p.p.523.

En la figura 1.1 se percibe el decremento de la pobreza previo a 1981 y su incremento paulatino de 1981 a 1994, indicado por las líneas CEPAL y Banco de México. La estimación HLB indicada en la figura 1.1 no se considera como parte de los comentarios pues resulta ser

⁶ Damian, Araceli y Boltvinik, Julio, *Evolución y características de la pobreza en México*, Comercio Exterior, Vol. 53, Núm. 6, junio 2003

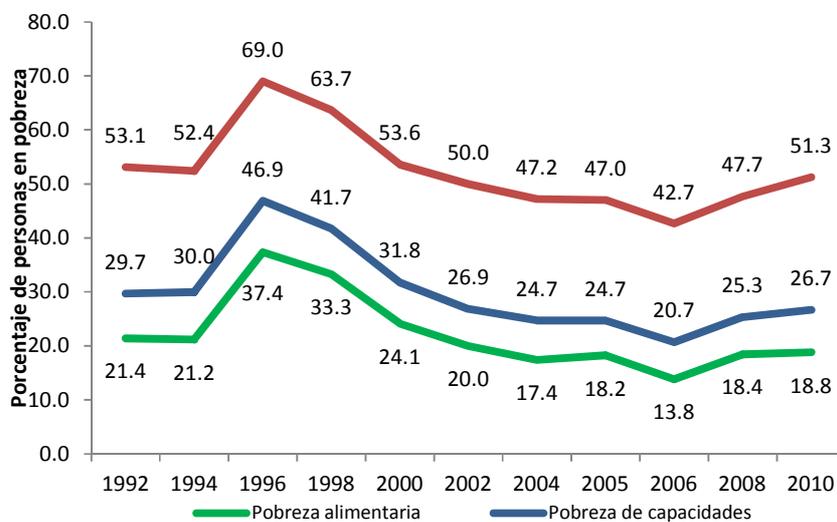
una estimación meramente de investigación, sin reconocimiento por alguna autoridad nacional o internacional.

Las variaciones presentadas para la figura 1.1 coinciden con las crisis económicas por las que atravesó el país y el incremento en el coste de los alimentos así como los demás bienes y servicios. Para el año 2000 disminuyó la proporción de pobres, congruente con la estabilización de la economía mexicana.

Al empatar la figura 1.1 y la figura 1.2 (pobreza de capacidades) en el año 1996, las proporciones de pobres se asemejan, haciéndolas medianamente comparables, considerando que las mediciones realizadas por CONEVAL y la Comisión Económica para América Latina (CEPAL) se presentan diferencias.

La pobreza en México disminuyó después de 1996 y las acciones gubernamentales, ya sea nacionales o de entidades internacionales, se han enfocado a minimizar el número de pobres, resultado que se visualiza en las gráficas y mediciones realizadas por CONEVAL.

Figura 1.2
Porcentaje de personas en pobreza por la dimensión de ingreso
1992 - 2010



Nota: las estimaciones de 2006, 2008 y 2010 utilizan los factores de expansión ajustados a los resultados definitivos del Censo de Población y Vivienda 2010, estimados por INEGI.

Fuente: Estimaciones del CONEVAL con base en las ENIGH de 1992 a 2010

1.1.1 Definición unidimensional de la pobreza

Previo a la promulgación de la Ley General de Desarrollo Social⁷, la medición de la pobreza en México se realizaba de acuerdo a la línea de bienestar. En 1993, el INEGI y la CEPAL publicaron el primer estudio de evolución de la pobreza en México, donde se establecen tres líneas de pobreza⁸. Las definiciones aún se conservan y forman parte del punto de comparación entre la información previa a la promulgación de la Ley.

Pobreza alimentaria

La pobreza alimentaria responde a la clasificación previa al decreto de la Ley General de Desarrollo Social, en la que se realizaban distinciones entre la población de acuerdo al ingreso, siendo así una medición unidimensional.

De acuerdo con el Glosario publicado por CONEVAL⁹, la pobreza alimentaria es:

Incapacidad para obtener una canasta básica alimentaria, aun si se hiciera uso de todo el ingreso disponible en el hogar para comprar sólo los bienes de dicha canasta.

Lo anterior implica un monto a considerar como la línea con la que determina ser o no pobre, considerando los ingresos para adquirir alimentos básicos.

Pobreza patrimonial

El término pobreza patrimonial también se encuentra asociado a la definición de pobreza en función del ingreso.

La pobreza patrimonial al igual que la definición de pobreza tiene múltiples definiciones; la utilizada en Plan Nacional de Desarrollo¹⁰ del gobierno para el período 2007 – 2012:

⁷ Ley General de Desarrollo Social, Ciudad de México, Diario Oficial de la Federación, 16 de junio 2012.

⁸ Damian, Araceli y Boltvinik, Julio, *Evolución y características de la pobreza en México*, Comercio Exterior, Vol. 53, Núm. 6, junio 2003

⁹ Consejo Nacional de Evaluación de la Política de Desarrollo social, *Glosario*, <http://web.coneval.gob.mx/Medicion/Paginas/Glosario.aspx>

¹⁰ *Plan nacional de desarrollo 2007 - 2012*, Ciudad de México, Gobierno de los Estados Unidos Mexicanos, mayo 2007, p.145.

Es la población que si bien puede cubrir sus necesidades mínimas de alimentación, educación y salud, cuenta con un ingreso per cápita que no le es suficiente para adquirir mínimos indispensables de vivienda, vestido, calzado y transporte para cada uno de los miembros del hogar.

Pobreza de capacidades

La pobreza de capacidades es un conjunto de servicios básicos, de salud y de educación de los que se carece y por tanto, la población que lo padece se encuentra vulnerable y con desventaja respecto a la población que posee estos servicios.

La definición establecida en el Plan Nacional de Desarrollo 2007 – 2012¹¹ considera a la población en situación de pobreza como:

La población que si bien puede cubrir sus necesidades mínimas de alimentación, cuenta con un ingreso per cápita insuficiente como para realizar las inversiones mínimamente aceptables en la educación y la salud de cada uno de los miembros del hogar.

Marginación social

La desigualdad social es la condición con la que las personas tienen acceso diferenciado a los recursos, servicios y posiciones que la sociedad valora. La marginación se encuentra asociada con las distintas posiciones de la estructura social y generalmente se enfoca en marcar diferencias. Por lo tanto, la desigualdad surge de la diferenciación social pues algunos papeles o posiciones sociales colocan a ciertas personas en condiciones de adquirir una mayor porción de bienes y servicios valorados.

Esta marginación afecta a algunos grupos sociales (indígenas, personas con discapacidad, personas de la tercera edad, entre otros) y generalmente estos grupos sociales reciben con mayor crudeza la carencia de oportunidades respecto al resto de la sociedad.

La marginación nunca se ha considerado como parte de la línea de pobreza, sin embargo al ojo del investigador es notorio que esta condición diferenciada debe ser estudiada para determinar el grado de afectación por tal situación.

¹¹ *Plan nacional de desarrollo 2007 - 2012*, Ciudad de México, Gobierno de los Estados Unidos Mexicanos, mayo 2007, p.145.

1.2 Los factores no económicos que determinan la pobreza

La pobreza es consecuencia de diversos factores, tanto económicos, políticos y sociales. El factor económico es influenciado por las políticas de los gobiernos al estimular o contraer el mercado, limitando así las posibilidades de crecimiento para la población.

Es importante considerar que la pobreza también posee un componente social y que existen factores que la perpetúan al mantener las condiciones. Phil Bartle¹² describe los siguientes factores como elementos que coadyuvan al origen de la pobreza:

- a) Ignorancia
- b) Enfermedad
- c) Apatía
- d) Corrupción
- e) Dependencia

Cada factor afecta a la sociedad de una forma distinta. Las afectaciones se presentan a continuación:

- *Ignorancia.* Este factor se asocia a la falta de conocimientos o de información relevante para mitigar la pobreza y la falta de información ocurre por: i) Carencia en la transmisión de la información al resto de la comunidad; o ii) Conservación del conocimiento por unos cuantos que desean aventajar al resto, limitando así la transmisión del conocimiento.

Cuando se menciona *información relevante* es preciso aclarar que ésta debe ser tal que contribuya con el desarrollo de la sociedad en la que se educa proporcionando medios de subsistencia y crecimiento en la comunidad. Conocimientos en agricultura, temporadas de lluvia para sembrar alimentos, tratamiento de animales domesticados, procesamiento de alimentos, entre otros, son ejemplos de educación que contribuye al desarrollo.

- *Enfermedad.* La enfermedad en una sociedad es un factor que conserva la pobreza. Si la sociedad se encuentra enferma en una proporción alta, el ausentismo es considerable y en consecuencia, la productividad baja y generando menos riqueza. Adicionalmente, una sociedad enferma eroga una mayor cantidad de recursos por conceptos médicos con alto costo, considerando que los tratamientos se enfocan a corregir más que a prevenir.

¹²Bartle, Phil trad. Sada Ma. Loudes, *Los cinco principales factores de la pobreza. Factores, no síntomas*, <http://cec.vcn.bc.ca/mpfc/modules/emp-povs.htm> (en español), Community Empowerment Collective. Consultado el 13 de julio, 2013..

Este factor se vincula con la carencia o adquisición de conocimientos de relevancia para la prevención de enfermedades en los aspectos básicos (lavado de manos previo a la ingesta de alimentos, higiene bucal, aseo diario, alimentación, etc).

- *Apatía*. Este factor describe la falta de interés de la sociedad respecto a ejecutar acciones o iniciativas para participar en la búsqueda de un bien colectivo para tener un beneficio o cambio individual. También suele vincularse a la sensación de impotencia para modificar las cosas y por lo tanto no se intenta ejecutar acción alguna que cambie de manera permanente su realidad.
- *Corrupción*. La corrupción se encuentra vinculada con la carencia de valores aprendidos en la sociedad en la que se presenta y al *efecto multiplicador* mencionado por los economistas. Éste efecto considera que una cantidad invertida o retirada en cuestión de materia social tiene un efecto, positivo o negativo según sea el caso, en una proporción mayor a la cantidad creada.

La corrupción y la apatía se vinculan para así crear un círculo vicioso sin fin perpetuando la formación de pobres por la carencia de deseos de cambio para la situación.

- *Dependencia*. La dependencia es el resultado de ser el extremo receptor de la caridad a largo plazo. La creencia que las personas requieren apoyo cuando se encuentran desvalidas y la autojustificación, contribuyen a perpetuar la condición de desvalido.

Entonces, la asistencia debe enfocarse en adiestrar a la comunidad creando capital y control del desarrollo propio. Esta afirmación entonces se vincula con los conocimientos que la comunidad debe adquirir. Un claro ejemplo de la modificación y eliminación de la dependencia es la redefinición de la Ley de Pobres de Inglaterra en 1834.

Todos los factores se encuentran correlacionados lo que provoca que uno derive en otro y simultáneamente existan afectaciones en el resto.

Al erradicar o disminuir estos factores se combate directamente a las consecuencias que afectan a la población en situación de pobreza, así como la reproducción de los pobres. Es importante mencionar que en el caso de México la población en situación de pobreza se mantiene constante a lo largo de las generaciones, es decir, los niños que fueron pobres, hoy continúan siendo pobres y procrean hijos pobres, lo cual resulta preocupante considerando que existe en la sociedad mexicana un factor adicional inexplorado que garantiza la continuidad de la pobreza.

Figura 1.3
Correlación entre los factores de la pobreza



Fuente: Los cinco principales factores de la pobreza. Factores, no síntomas. CEC.

1.3 Análisis cuantitativo de la pobreza

Analizar y medir la pobreza es relevante para administrar y garantizar la cobertura de los programas sociales puestos en marcha por el Gobierno Federal, Estatal o Municipal, evaluando el impacto de los recursos del erario público destinados a disminuir o mitigar los efectos de la pobreza en sus distintos niveles.

En el informe presentado por el CONEVAL¹³ se describe la reducción de las carencias sociales como parte de la ejecución de programas sociales; adicionalmente, destaca el incremento del número de personas con acceso a la alimentación. Respecto al ingreso y en el contexto de la crisis económica, el ingreso monetario real percibido por los hogares mexicanos se redujo considerablemente.

En consecuencia, la población en pobreza en el país aumentó de 44.5% a 46.2%, equivalente a 48.8 millones de personas entre 2008 y 2010. En el mismo período, la población en pobreza extrema disminuyó de 10.6% a 10.4%. Para el período 2008 -2010, a nivel nacional se obtuvieron los siguientes resultados:

¹³ Consejo Nacional de Evaluación de la Política de Desarrollo social, *Informe de la Evaluación de la Política de Desarrollo Social en México 2010*, México, D.F., CONEVAL, 2012.

Cuadro 1.2
Medición de la Pobreza, Estados Unidos Mexicanos, 2010

Indicadores	Porcentaje		Millones de personas		Carencias promedio	
	2008	2010	2008	2010	2008	2010
Pobreza						
Población en situación de pobreza	44.5	46.2	48.8	52.0	2.7	2.5
Población con pobreza moderada	33.9	35.8	37.2	40.3	2.3	2.1
Población en situación de pobreza extrema	10.6	10.4	11.7	11.7	3.9	3.7
Población vulnerable por carencias sociales	33.0	28.7	36.2	32.3	2.0	1.9
Población vulnerable por ingresos	4.5	5.8	4.9	6.5	0.0	0.0
Población no pobre y no vulnerable	18.0	19.3	19.7	21.8	0.0	0.0
Privación social						
Población con al menos una carencia social	77.5	74.9	85.0	84.3	2.4	2.3
Población con al menos tres carencias sociales	31.1	26.6	34.1	29.9	3.7	3.6
Indicadores de carencia social						
Rezago educativo	21.9	20.6	24.1	23.2	3.2	3.0
Carencia por acceso a los servicios de salud	40.8	31.8	44.8	35.8	2.9	2.8
Carencia por acceso a la seguridad social	65.0	60.7	71.3	68.3	2.6	2.5
Carencia por calidad y espacios de vivienda	17.7	15.2	19.4	17.1	3.6	3.5
Carencia por acceso a los servicios básicos en la vivienda	19.2	16.5	21.1	18.5	3.5	3.3
Carencia por acceso a la alimentación	21.7	24.9	23.8	28.0	3.3	3.0
Bienestar						
Población con un ingreso inferior a la línea de bienestar mínimo	16.7	19.4	18.4	21.8	3.0	2.7
Población con un ingreso inferior a la línea de bienestar	49.0	52.0	53.7	58.5	2.5	2.2

Fuente: Estimaciones del CONEVAL con base en el MCS-ENIGH 2008 y 2010.

Con la información contenida en el cuadro 1.2, es fácil verificar que la población con ingresos inferiores a la línea de bienestar se incrementó en 2010 respecto a 2008, consecuencia de la crisis económica mundial vivida en los años citados.

Para la población mexicana, el acceso a la seguridad social habitualmente se consigue mediante el empleo formal pese a las distintas maneras de acceder a ellas voluntariamente. Al disminuirse el empleo formal entre 2008 y 2010, consecuencia de la crisis internacional que también afectó a México, es evidente que la carencia social con la mayor proporción es la seguridad social.

El aumento en los precios de los alimentos consecuentemente afecta el acceso a la alimentación y traslada a los más desvalidos a carecer de la correcta alimentación, fenómeno que se confirma con el incremento del 3.2% de 2008 al 2010.

1.3.1 Para las entidades federativas con mayor proporción de personas en situación de pobreza en México

Al evaluar a la población en situación de pobreza, destacan las siguientes entidades por ser las poseedoras de la proporción de pobres más alta del país: Chiapas, Guerrero, Oaxaca, Tlaxcala y Zacatecas. Los estados con el mayor número de personas pobres son Chiapas, México y Veracruz. A continuación se analiza la situación para cada uno de las entidades citadas.

Chiapas

En el estado de Chiapas es donde se concentra la mayor proporción de población en situación de pobreza con 78.4%. La población pobre se incrementó en 1.4% de 2008 a 2010. Esta proporción se ve influenciada por la carencia de ingresos mínimos para la adquisición de alimentos y bienes. Adicionalmente es notorio que el 80.9 % de la población tiene un ingreso inferior a la línea de bienestar a 2010.

Cuadro 1.3
Medición de la Pobreza, Chiapas, 2010

Incidencia, número de personas y carencias promedio en los indicadores de pobreza, 2008-2010

Indicadores	Porcentaje		Miles de personas		No. Carencias promedio	
	2008	2010	2008	2010	2008	2010
Pobreza						
Población en situación de pobreza	77.0	78.4	3,573.4	3,777.7	3.2	2.9
Población en situación de pobreza moderada	41.4	45.6	1,920.9	2,197.1	2.5	2.2
Población en situación de pobreza extrema	35.6	32.8	1,652.5	1,580.6	3.9	3.8
Población vulnerable por carencias sociales	16.1	13.1	746.5	633.1	2.1	2.1
Población vulnerable por ingresos	1.5	2.5	70.8	119.2	0.0	0.0
Población no pobre y no vulnerable	5.4	6.0	252.4	289.7	0.0	0.0
Privación social						
Población con al menos una carencia social	93.0	91.5	4,319.9	4,410.8	3.0	2.8
Población con al menos tres carencias sociales	57.2	48.7	2,656.7	2,349.0	3.8	3.7
Indicadores de carencia social						
Rezago educativo	38.0	35.0	1,763.1	1,687.6	3.7	3.4
Carencia por acceso a los servicios de salud	52.2	36.5	2,423.1	1,758.3	3.5	3.5
Carencia por acceso a la seguridad social	85.4	82.4	3,966.3	3,970.0	3.1	2.9
Carencia por calidad y espacios de vivienda	38.4	33.2	1,780.8	1,602.5	3.8	3.7
Carencia por acceso a los servicios básicos en la vivienda	36.3	34.1	1,685.1	1,642.1	3.9	3.6
Carencia por acceso a la alimentación	26.2	30.3	1,216.8	1,460.9	3.9	3.6

Incidencia, número de personas y carencias promedio en los indicadores de pobreza, 2008-2010

Indicadores	Porcentaje		Miles de personas		No. Carencias promedio	
	2008	2010	2008	2010	2008	2010
Bienestar						
Población con un ingreso inferior a la línea de bienestar mínimo	48.2	50.9	2,239.2	2,452.8	3.3	3.1
Población con un ingreso inferior a la línea de bienestar	78.5	80.9	3,644.2	3,896.9	3.1	2.8

Fuente: Estimaciones del CONEVAL con base en el MCS-ENIGH 2008 y 2010.

Guerrero

En Guerrero la población en situación de pobreza tuvo un decremento entre 2008 y 2010 consecuencia de algunos programas sociales, sin embargo, el 90.5% de la población tiene al menos una carencia social, lo cual resulta relevante.

La población con acceso a la seguridad social disminuyó en el período de análisis, lo que denota una disminución en la cantidad de empleos formales mantenidos y creados de 2008 a 2010.

Cuadro 1.4
Medición de la Pobreza, Guerrero, 2010

Incidencia, número de personas y carencias promedio en los indicadores de pobreza, 2008-2010

Indicadores	Porcentaje		Miles de personas		No. Carencias promedio	
	2008	2010	2008	2010	2008	2010
Pobreza						
Población en situación de pobreza	68.4	67.4	2,282.3	2,286.4	3.4	3.2
Población en situación de pobreza moderada	37.0	38.6	1,236.3	1,309.2	2.8	2.6
Población en situación de pobreza extrema	31.3	28.8	1,046.0	977.2	4.1	4.0
Población vulnerable por carencias sociales	23.0	23.1	768.2	782.4	2.5	2.4
Población vulnerable por ingresos	2.0	2.1	68.1	71.6	0.0	0.0
Población no pobre y no vulnerable	6.6	7.4	218.5	250.1	0.0	0.0
Privación social						
Población con al menos una carencia social	91.4	90.5	3,050.5	3,068.8	3.2	3.0
Población con al menos tres carencias sociales	60.9	55.2	2,033.3	1,872.6	4.0	3.8
Indicadores de carencia social						
Rezago educativo	28.5	28.3	951.4	958.4	4.0	3.7
Carencia por acceso a los servicios de salud	57.7	39.6	1,925.4	1,342.0	3.7	3.6
Carencia por acceso a la seguridad social	81.2	78.4	2,708.7	2,659.4	3.4	3.2
Carencia por calidad y espacios de la vivienda	44.7	40.6	1,491.3	1,374.9	4.0	3.8
Carencia por acceso a los servicios básicos	46.2	40.1	1,540.6	1,358.7	4.0	3.8

Incidencia, número de personas y carencias promedio en los indicadores de pobreza, 2008-2010

Indicadores	Porcentaje		Miles de personas		No. Carencias promedio	
	2008	2010	2008	2010	2008	2010
en la vivienda						
Carencia por acceso a la alimentación	34.0	42.6	1,134.0	1,443.3	4.1	3.6
Bienestar						
Población con un ingreso inferior a la línea de bienestar mínimo	38.9	38.5	1,298.1	1,305.6	3.7	3.4
Población con un ingreso inferior a la línea de bienestar	70.4	69.5	2,350.5	2,358.0	3.3	3.1

Fuente: estimaciones del CONEVAL con base en el MCS-ENIGH 2008 y 2010.

Oaxaca

El caso de Oaxaca es el escenario perfecto para ejemplificar los efectos de la crisis económica del 2008 y la influencia de los programas sociales llevados a cabo entre 2008 y 2010.

Al 2010 el 67.2% de su población se encontraba en situación de pobreza y la población con al menos una privación social se mantuvo, sin embargo la población con un ingreso menor a la línea de bienestar se incrementó a quedar en un 68.6%, con un 5.4% de incremento.

Adicionalmente en la entidad se presenta una mejora en la población con acceso a los servicios de salud y médicos, consecuencia de los programas sociales.

**Cuadro 1.5
Medición de la Pobreza, Oaxaca, 2010**

Incidencia, número de personas y carencias promedio en los indicadores de pobreza, 2008-2010

Indicadores	Porcentaje		Miles de personas		No. Carencias promedio	
	2008	2010	2008	2010	2008	2010
Pobreza						
Población en situación de pobreza	61.8	67.2	2,310.4	2,557.5	3.5	3.0
Población en situación de pobreza moderada	34.1	40.5	1,274.8	1,543.9	3.0	2.4
Población en situación de pobreza extrema	27.7	26.6	1,035.6	1,013.5	4.2	3.9
Población vulnerable por carencias sociales	27.4	22.1	1,023.2	841.5	2.4	2.2
Población vulnerable por ingresos	1.4	1.4	53.1	54.5	0.0	0.0
Población no pobre y no vulnerable	9.4	9.3	351.5	354.3	0.0	0.0
Privación social						
Población con al menos una carencia social	89.2	89.3	3,333.6	3,399.0	3.2	2.8
Población con al menos tres carencias sociales	58.4	49.4	2,184.1	1,880.5	4.0	3.8
Indicadores de carencia social						
Rezago educativo	30.6	30.3	1,144.7	1,152.1	4.0	3.7
Carencia por acceso a los servicios de salud	56.2	39.9	2,102.0	1,518.1	3.7	3.5

Incidencia, número de personas y carencias promedio en los indicadores de pobreza, 2008-2010

Indicadores	Porcentaje		Miles de personas		No. Carencias promedio	
	2008	2010	2008	2010	2008	2010
Carencia por acceso a la seguridad social	80.4	79.7	3,006.1	3,033.8	3.3	2.9
Carencia por calidad y espacios de vivienda	38.2	34.1	1,428.6	1,299.4	4.1	3.7
Carencia por acceso a los servicios básicos en la vivienda	48.5	41.4	1,812.0	1,575.8	3.9	3.6
Carencia por acceso a la alimentación	28.6	26.6	1,068.2	1,014.0	4.2	3.8
Bienestar						
Población con un ingreso inferior a la línea de bienestar mínimo	32.9	36.8	1,231.6	1,402.0	3.8	3.3
Población con un ingreso inferior a la línea de bienestar	63.2	68.6	2,363.5	2,612.0	3.4	3.0

Fuente: Estimaciones del CONEVAL con base en el MCS-ENIGH 2008 y 2010.

Tlaxcala

En el caso de Tlaxcala, es evidente que la población que se encuentra en situación de pobreza lo está porque los ingresos son insuficientes, pues los índices de carencia social en todos los rubros tuvieron una mejoría en 2010 respecto a 2008.

La afectación fue más fuerte entre los vulnerables por ingreso, pues aumentó el número de personas cuyo ingreso se mermó, alcanzando la línea de bienestar.

**Cuadro 1.6
Medición de la Pobreza, Tlaxcala, 2010**

Incidencia, número de personas y carencias promedio en los indicadores de pobreza, 2008-2010

Indicadores	Porcentaje		Miles de personas		No. Carencias promedio	
	2008	2010	2008	2010	2008	2010
Pobreza						
Población en situación de pobreza	59.8	60.4	677.5	710.8	2.3	2.1
Población en situación de pobreza moderada	50.9	51.2	577.2	602.2	2.1	1.9
Población en situación de pobreza extrema	8.9	9.2	100.3	108.6	3.6	3.4
Población vulnerable por carencias sociales	24.2	19.9	274.7	234.2	2.0	1.8
Población vulnerable por ingresos	5.4	7.2	61.1	84.2	0.0	0.0
Población no pobre y no vulnerable	10.6	12.5	120.0	147.0	0.0	0.0
Privación social						
Población con al menos una carencia social	84.0	80.3	952.3	945.1	2.2	2.1
Población con al menos tres carencias sociales	28.2	21.8	319.9	256.3	3.4	3.4
Indicadores de carencia social						
Rezago educativo	17.7	15.6	200.8	183.3	3.0	2.8
Carencia por acceso a los servicios de salud	46.7	35.1	529.7	412.5	2.7	2.6
Carencia por acceso a la seguridad social	74.9	71.1	849.1	836.6	2.3	2.1

Incidencia, número de personas y carencias promedio en los indicadores de pobreza, 2008-2010

Indicadores	Porcentaje		Miles de personas		No. Carencias promedio	
	2008	2010	2008	2010	2008	2010
Carencia por calidad y espacios de la vivienda	13.4	11.8	151.6	139.1	3.3	3.2
Carencia por acceso a los servicios básicos en la vivienda	9.4	6.9	106.9	81.4	3.5	3.5
Carencia por acceso a la alimentación	24.8	24.3	281.1	285.4	3.0	2.8
Bienestar						
Población con un ingreso inferior a la línea de bienestar mínimo	21.7	26.7	246.4	314.6	2.4	2.2
Población con un ingreso inferior a la línea de bienestar	65.2	67.6	738.6	795.0	2.1	1.9

Fuente: Estimaciones del CONEVAL con base en el MCS-ENIGH 2008 y 2010.

Zacatecas

En Zacatecas el movimiento de la población general tuvo un matiz interesante, pues los rubros asociados a ingresos se mermaron. Los indicadores de carencias sociales se modificaron aunque la carencia por alimentación incrementó 5.1% en dos años. El resultado es una mayor concentración de personas en pobreza moderada.

Cuadro 1.7
Medición de la Pobreza, Zacatecas, 2010

Incidencia, número de personas y carencias promedio en los indicadores de pobreza, 2008-2010

Indicadores	Porcentaje		Miles de personas		No. Carencias promedio	
	2008	2010	2008	2010	2008	2010
Pobreza						
Población en situación de pobreza	50.4	60.2	740.3	899.0	2.3	2.1
Población en situación de pobreza moderada	40.9	49.8	600.6	743.3	2.0	1.9
Población en situación de pobreza extrema	9.5	10.4	139.7	155.7	3.6	3.5
Población vulnerable por carencias sociales	28.1	18.7	412.0	279.0	1.9	1.8
Población vulnerable por ingresos	5.8	6.8	84.6	101.4	0.0	0.0
Población no pobre y no vulnerable	15.7	14.3	230.6	214.1	0.0	0.0
Privación social						
Población con al menos una carencia social	78.5	78.9	1,152.3	1,178.0	2.2	2.1
Población con al menos tres carencias sociales	25.1	22.8	368.5	340.5	3.5	3.4
Indicadores de carencia social						
Rezago educativo	24.5	22.8	360.0	340.3	2.8	2.6
Carencia por acceso a los servicios de salud	33.5	27.0	491.9	402.7	2.8	2.7
Carencia por acceso a la seguridad social	67.8	66.4	995.6	991.8	2.3	2.2
Carencia por calidad y espacios de la vivienda	9.6	5.9	140.2	87.3	3.5	3.6

Incidencia, número de personas y carencias promedio en los indicadores de pobreza, 2008-2010

Indicadores	Porcentaje		Miles de personas		No. Carencias promedio	
	2008	2010	2008	2010	2008	2010
Carencia por acceso a los servicios básicos en la vivienda	14.2	16.1	208.6	240.3	3.2	3.0
Carencia por acceso a la alimentación	19.7	24.8	288.8	370.3	3.0	2.7
Bienestar						
Población con un ingreso inferior a la línea de bienestar mínimo	22.3	29.5	327.5	440.7	2.4	2.2
Población con un ingreso inferior a la línea de bienestar	56.2	67.0	824.9	1,000.5	2.0	1.9

Fuente: Estimaciones del CONEVAL con base en el MCS-ENIGH 2008 y 2010.

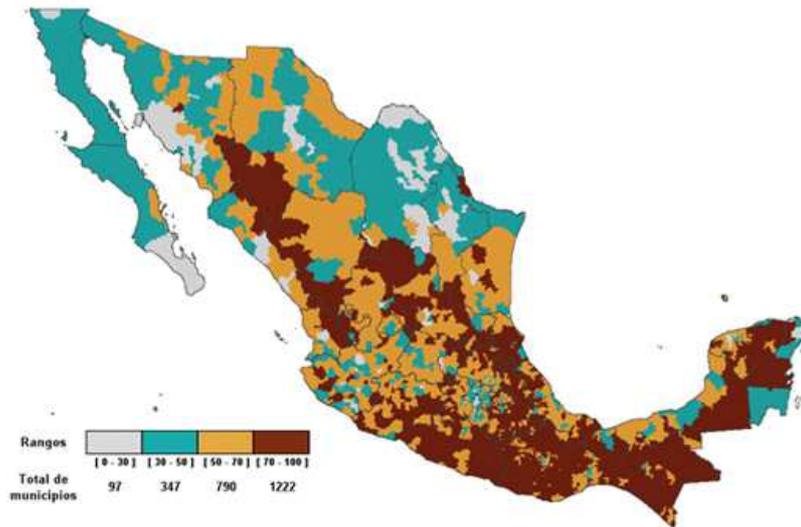
1.3.2 Por zona de residencia (rural - urbano)

En cada ciudad próspera existe un cinturón o zona de miseria que la rodea. Éste resulta ser un problema del gobierno, pues al no existir oportunidades de crecimiento en las pequeñas comunidades o zonas rurales, la población se moviliza a las grandes urbes, creando asentamientos irregulares y fuera de norma. En consecuencia, el hacinamiento en zonas alejadas de la prosperidad de la ciudad agudiza la pobreza por la carencia de servicios y por el perfil mismo de la gente que habita ahí.

Es complicada la situación del Estado pues estos asentamientos no proveen de impuestos para su manutención y a su vez, generan costos exorbitantes para la introducción de los servicios que debe prestarse (agua, luz, drenaje, planteles educativos, seguridad y servicios médicos).

El panorama a nivel municipal se visualiza con facilidad en la figura 1.2. Los municipios con mayor concentración de pobres, resaltados en color marrón, se encuentran en el sur del país o bien, en las grandes urbes.

Figura 1.4
Medición de pobreza 2010 por municipio. Porcentaje de la población en situación de pobreza por municipio, México, 2010



Fuente: Estimaciones del CONEVAL con base en el MCS-ENIGH 2010 y la muestra del Censo de Población y Vivienda 2010.

Los resultados respecto a la pobreza municipal en el estudio realizado por el CONEVAL para 2010 se presentan en dos grupos, los municipios con mayor proporción de pobreza, detallados en el cuadro 1.8.

Cuadro 1.8
Municipios con mayores porcentajes de población en situación de pobreza

Municipio		Estado	Porcentaje de la población
Mayor porcentaje de población en pobreza	San Juan Tepeuxila	Oaxaca	97.4
	Aldama	Chiapas	97.3
	San Juan Cancuc	Chiapas	97.3
	Mixtla de Altamirano	Veracruz	97
	Chalchihuitán	Chiapas	96.8

Fuente: Elaboración propia con estimaciones del CONEVAL con base en el MCS-ENIGH 2010.

Los municipios con mayor proporción de pobreza son localidades rurales o que se encuentran alejadas de las cabeceras municipales, pues el acceso a servicios médicos y demás servicios proporcionados por el Estado se limitan por la distancia a recorrer para hacer uso de ellos.

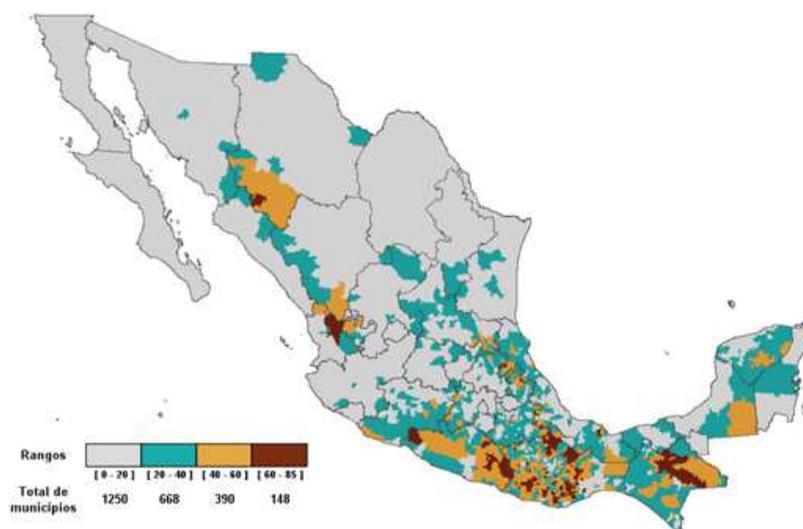
En el cuadro 1.9 se muestran los municipios con mayor cantidad de pobres. Éstos corresponden a zonas urbanas, confirmando así que la pobreza se acentúa en las zonas aledañas a las grandes urbes, donde los servicios no se proveen por la falta de oportunidades y la migración de la gente del campo a la ciudad.

Cuadro 1.9
Municipios con mayor cantidad de gente en situación de pobreza

Municipio		Estado	Número de personas
Mayor número de personas en pobreza	Puebla	Puebla	723 154
	Iztapalapa	Distrito Federal	727 128
	Ecatepec de Morelos	México	723 559
	León	Guanajuato	600 145
	Tijuana	Baja California	525 769
	Juárez	Chihuahua	494 726
	Nezahualcóyotl	México	462 405
	Toluca	México	407 691
	Acapulco de Juárez	Guerrero	405 499
	Gustavo A. Madero	Distrito Federal	356 328

Fuente: Elaboración propia con estimaciones del CONEVAL con base en el MCS-ENIGH 2010.

Figura 1.5
Porcentaje de la población en situación de pobreza extrema por municipio, México, 2010



Fuente: Estimaciones del CONEVAL con base en el MCS-ENIGH 2010 y la muestra del Censo de Población y Vivienda 2010.

1.4 Mediciones de la pobreza en México

Para comprender la complejidad de la medición de la pobreza es importante abordar la trayectoria y evolución de las mediciones de la pobreza en México.

En este apartado se menciona el antes y el después de la promulgación de la Ley General de Desarrollo Social, siendo ésta un parte aguas para obtener la medida de la pobreza a nivel nacional, consecuencia de los compromisos adquiridos con las instituciones internacionales para disminuir la cantidad de personas en situación de pobreza y de pobreza extrema.

1.4.1 Índice de pobreza de acuerdo a ingreso anterior a 2004

Por muchas décadas no hubo medición oficial de pobreza en México pues las políticas no contemplaban una medida, además que la información existente se encontraba dispersa, pues fue hasta 1983 cuando se constituyó el Instituto Nacional de Estadística y Geografía (INEGI). En consecuencia, las acciones sociales no se encontraban focalizadas para atender a la población vulnerable.

En el año 2000 la Organización de las Naciones Unidas, en colaboración con los 189 países miembros, establecieron los Objetivos del Milenio. Los ocho objetivos a cumplir son:

- 1) Erradicar la pobreza extrema y el hambre.
- 2) Lograr la enseñanza primaria universal.
- 3) Promover la igualdad entre los sexos y el empoderamiento de la mujer.
- 4) Reducir la mortalidad de los niños menores de 5 años.
- 5) Mejorar la salud materna.
- 6) Combatir el VIH/SIDA, la malaria y otras enfermedades.
- 7) Garantizar la sostenibilidad del medio ambiente.
- 8) Fomentar una alianza mundial para el desarrollo.

Dichos objetivos deben cumplirse al 2015 y, como cualquier objetivo deben ser medibles, por lo tanto, entre 2001 y 2002, la SEDESOL elaboró la primera medición oficial de pobreza a nivel nacional.

La medida era simple: quien tuviera ingresos por debajo de un monto establecido (línea de pobreza), era considerado pobre. Había tres líneas de pobreza: alimentaria, de capacidades y de patrimonio.

Dado el requerimiento para realizar una medición y comparación posterior, el INEGI realizó ajustes a sus cuestionarios, bases de datos y demás fuentes de información proporcionando

así los datos necesarios para estimar los indicadores de pobreza, previo al decreto de la Ley General de Desarrollo. Sin embargo, por la carencia de información, este proyecto no se alcanzó en su totalidad y aún a 2012 la información y estructuras de la información se encuentra en mejora continua.

1.4.2 Metodología multidimensional

La metodología multidimensional se estableció para México en 2006, con el decreto de la Ley General de Desarrollo Social. Dado que en ese momento se contaba únicamente con el lineamiento, fue necesario crear:

- a) El CONEVAL y su estructura para funcionar.
- b) Elaboración de un marco perfectamente definido para establecer los parámetros a medir y la metodología de medición.
- c) Estadísticas consistentes y congruentes con los marcos y modelos elaborados.

Hasta el 2008 se realizó una medición de la pobreza multidimensional, la cual considera todos los indicadores estipulados en la Ley.

Para 2010 y 2012, se realiza una nueva medición de la pobreza a nivel nacional presentando información comparable para monitorear el impacto que cada uno de los programas sociales tuvieron en la población. Asimismo, es asequible la elaboración de un plan a largo plazo para la mitigación de la pobreza.

1.5 Políticas públicas para mitigar la pobreza en el país

Desde mediados del siglo XX, en México se ha considerado la necesidad de apoyar a los grupos más vulnerables y acercar bienes y servicios que coadyuven a su desarrollo. En ese sentido, cada sexenio considera la creación y seguimiento de diversos planes de apoyo.

En este apartado se presentarán los programas de desarrollo social vigentes a la fecha de elaboración del proyecto.

1.5.1 Plan Nacional de Desarrollo (2007 – 2012)

En el Plan Nacional de Desarrollo elaborado por el Ejecutivo Federal al 2007, se consideró la *Igualdad de Oportunidades* como tercer eje prioritario para la administración. La superación de la pobreza consideró a su vez la situación al 2007 y se establecieron tres objetivos y sus estrategias para disminuirla.

Objetivo 1: Reducir significativamente el número de mexicanos en condiciones de pobreza con políticas públicas que superen un enfoque asistencialista, de modo que las personas puedan adquirir capacidades y generar oportunidades de trabajo.

- Fortalecer los programas existentes de superación de la pobreza, ampliándolos para incluir otras vertientes de apoyo social y asegurar que lleguen a la población que realmente los necesita.
- Impulsar prioritariamente el desarrollo de los municipios de mayor marginación, a través de una efectiva focalización de recursos y coordinación de esfuerzos entre el Gobierno Federal y los demás órdenes de gobierno.
- Asegurar que los mexicanos en situación de pobreza resuelvan sus necesidades de alimentación y vivienda digna, con pleno acceso a servicios básicos y a una educación y salud de calidad.
- Mejorar los procesos de planeación, operación y evaluación de los programas para la superación de la pobreza, incluyendo la elaboración de un padrón único de beneficiarios.
- Fortalecer y ampliar los programas e instrumentos del Gobierno Federal para el combate de la pobreza en las ciudades del país.
- Promover la participación pobreza.

Objetivo 2: Apoyar a la población más pobre a elevar sus ingresos y a mejorar su calidad de vida, impulsando y apoyando la generación de proyectos productivos.

- Apoyar el arranque y la operación de proyectos productivos familiares y de grupos comunitarios mediante asesoría y programas de microfinanciamiento, en el campo y las ciudades.
- Ampliar la cobertura y mejorar la calidad de las vías y medios de comunicación y de transporte para conectar a las regiones menos desarrolladas del país.
- Promover proyectos de ecoturismo, turismo de aventura y turismo cultural en las zonas rurales para que puedan aprovechar sus ventajas comparativas en cuanto a riqueza cultural y natural y hagan de ésta una actividad que detone su desarrollo económico y social.
- Reorientar y fortalecer los programas de las instituciones públicas del sector agropecuario para detonar el desarrollo de actividades económicas en el campo.
- Empezar acciones para propiciar el empleo en zonas donde se genera la expulsión de personas, procurando convertirlas en receptoras de inversión.

Objetivo 3: Lograr un patrón territorial nacional que frene la expansión desordenada de las ciudades, provea suelo apto para el desarrollo urbano y facilite el acceso a servicios y equipamientos en comunidades tanto urbanas como rurales.

- Promover el ordenamiento territorial, la certeza jurídica en la tenencia de la tierra y la seguridad pública en zonas marginadas de las ciudades.
- Impulsar el ordenamiento territorial nacional y el desarrollo regional a través de acciones coordinadas entre los tres órdenes de gobierno y concertadas con la sociedad civil.
- Prevenir y atender los riesgos naturales.
- Fortalecer el marco institucional federal en materia de desarrollo urbano creando los instrumentos financieros técnicos y normativos que requiere la problemática actual de nuestras ciudades.

1.5.2 Plan Visión 2030

El Plan Visión 2030 va en concordancia con el Plan Nacional de Desarrollo 2007 – 2012. En éste, se consideran metas concretas para disminuir la pobreza, abordando los indicadores de pobreza con el objetivo de disminuir la pobreza extrema, mejorando la calidad de vida de los mexicanos.

En materia de pobreza extrema, las metas para 2030 son:

- a) Ningún mexicano en pobreza extrema.
- b) Disminuir la desigualdad entre entidades para que ningún estado se encuentre a más de 10 puntos porcentuales por arriba del promedio nacional.
- c) Toda la población debe contar con el servicio de agua potable.
- d) Elevar el nivel promedio de 15 años de escolaridad.
- e) Alcanzar 500 puntos en la evaluación del *Program for International Students Assessment*.

1.5.3 Programa de apoyo alimentario en zonas de atención prioritaria

El objetivo general del Programa es contribuir al establecimiento de condiciones de igualdad de oportunidades para un desarrollo pleno de los hogares en las localidades más dispersas del país, proporcionándoles las herramientas que les permitan mejorar su alimentación y nutrición y fomentar el acceso a los beneficios de otros programas y servicios públicos.

La población objetivo de este programa son los hogares ubicados en:

- Localidades que no forman parte del universo de atención y que nunca hayan sido acreedores al programa Oportunidades o DICONSA;
- Localidades de alta o muy alta marginación, o
- Estar en condiciones de pobreza de capacidades conforme a los criterios establecidos por el CONEVAL.

Los tipos de apoyo que se reciben en el programa de apoyo alimentario en zonas de atención prioritaria son:

- Dinero en efectivo para mejorar la alimentación de los hogares objetivo.
- La entrega de complementos nutricionales, de acuerdo al número de niños y niñas en el hogar de entre 6 meses y 6 años, y al número de mujeres embarazadas o en periodo de lactancia, y
- Orientación y promoción social para el desarrollo del hogar y la organización comunitaria.

1.5.4 Programa de abasto social de leche

El Programa de Abasto Social de Leche inició en 1944, y ha tenido distintas reestructuras para convertirse en LICONSA, a partir de 1994, cuando el programa se resectorizó, de la Secretaría de Comercio y Fomento Industrial a la Secretaría de Desarrollo Social.

LICONSA está constituida como una empresa de participación estatal mayoritaria que trabaja con el propósito de mejorar los niveles de nutrición de millones de mexicanos, contribuyendo así a su incorporación al desarrollo del país, mediante el apoyo a la alimentación de familias en condiciones de pobreza patrimonial, con leche a precio subsidiado, fundamentalmente para niños de hasta 12 años. La población objetivo del programa:

- Niñas y niños de 6 meses a 12 años y niñas hasta los 15 años de edad.
- Mujeres de entre 45 y 59 años y en periodo de gestación y lactancia.
- Enfermos crónicos y personas con discapacidad mayores de 12 años y
- Adultos de 60 y más años.

1.5.5 Programa de desarrollo Oportunidades

El programa Oportunidades se enfoca a atender a los hogares cuyas condiciones socioeconómicas y de ingreso impiden desarrollar las capacidades de sus integrantes en

materia de alimentación, salud y educación, de acuerdo con las Reglas de operación que el mismo programa establece.

Las familias elegibles para ingresar o reingresar al Programa son aquellas cuyo ingreso mensual per cápita estimado es menor a la Línea de Bienestar Mínimo.

El programa considera los siguientes apoyos:

1. Recursos para mujeres, madres de familia, para el ingreso familiar y una mejor alimentación.
2. Becas para niños y jóvenes, a partir de tercero de primaria y hasta el último grado de educación media superior.
3. Apoyo monetario a familias beneficiarias con hijos de 0 a 9 años para fortalecer su desarrollo.
4. Fondo de ahorro para jóvenes que concluyen su Educación Media Superior.
5. Apoyo para útiles escolares.
6. Paquete de servicios médicos y sesiones educativas para la salud.
7. Suplementos alimenticios a niños y niñas entre 6 y 23 meses, y con desnutrición entre los 2 y 5 años. También a las mujeres embarazadas o en periodo de lactancia.
8. Apoyo adicional por cada adulto mayor integrante de las familias beneficiarias, que no reciba recursos del Programa de Pensión para Adultos Mayores de la SEDESOL.

1.5.6 Comisión Nacional de Derechos Humanos

La Comisión Nacional de Derechos Humanos considera dentro de sus funciones el seguimiento y monitoreo de comunidades o personas que se encuentran en situación de pobreza, sin embargo, la atención se prioriza de acuerdo a la severidad de los temas.

El seguimiento permite considerar un apoyo dentro de esta institución que garantiza que bajo cualquier circunstancia, las personas que se encuentran vulnerables ante otros peligros perseguidos por la CNDH, alertarán a la sociedad y al Estado para evitar que nuevos casos se presenten.

Capítulo 2

Modelo lineal generalizado para la estimación de la pobreza extrema

Introducción

Los Modelos Lineales Simples, también conocidos como modelos de regresión lineal o el principio de mínimos cuadrados, se utilizan cuando la variable dependiente se relaciona linealmente con la variable independiente. Sin embargo, ciertos conjuntos de datos asociados a estudios biométricos, de fenómenos sociales o estudios físicos no cumplen con la relación lineal los Modelos Lineales Simples o Múltiples.

La respuesta a tal planteamiento son los Modelos Lineales Generalizados (MLG) pues son la ampliación o generalización de los modelos lineales pues se enfocan en patrones de variación, subsanando la necesidad de aproximar a datos que no se ajustan a los Modelos Lineales Simples.

En el primer apartado de este capítulo se plantean la documentación del uso de un modelo lineal generalizado para la medición de la pobreza.

En el segundo apartado se profundiza en la teoría de los Modelos Lineales Generalizados, haciendo especial énfasis en aquellos especificados para datos cuya respuesta sea dicotómica para analizar la proporción de la población en pobreza extrema.

En el tercer apartado se describe la metodología que sustenta el uso de los Modelos Lineales Generalizados en la estimación de los índices de pobreza extrema abordado en el proyecto.

2.1 Uso de un modelo lineal generalizado para estimar la proporción de la pobreza extrema

La Comisión Económica para América Latina y el Caribe (CEPAL) y la Oficina Regional del Fondo de las Naciones Unidas para la Infancia para América Latina y el Caribe (UNICEF-TACRO), desde 2008 adoptaron la metodología para la medición de la pobreza infantil desarrollada en 2003 por UNICEF, la Universidad de Bristol y la London School of Economics (metodología de medición de privaciones múltiples conocida como Indicadores de Bristol).

Esta metodología considera el enfoque multidimensional, cuyo resultado es *Pobreza infantil en América Latina y el Caribe*¹⁴; guía que presenta una opción para estimar la desnutrición infantil con medidas de encuestas en hogares; lo anterior guarda una similitud con la estimación de pobreza de este proyecto, al estimar o clasificar a una persona como pobre de acuerdo a ciertas variables categóricas y posteriormente estimar la probabilidad de ser pobre extremo.

2.2 Notación y nomenclatura

Con el objeto de homologar la nomenclatura matemática presentada en el desarrollo del proyecto, se especifica en ésta sección la notación a utilizar, con su respectivo significado.

Términos asociados al Álgebra Lineal

Los vectores serán representados con las letras griegas $\alpha, \beta, \gamma, \eta$, tales que

$$\beta = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} \quad (1)$$

Sea entonces $\beta^T = (\beta_1 \quad \dots \quad \beta_p)$, el vector transpuesto de β .

Las matrices se representarán con las letras mayúsculas del alfabeto griego:

$$X = \begin{bmatrix} x_{11} & \dots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{p1} & \dots & x_{pn} \end{bmatrix} \quad (2)$$

Asociados a la Probabilidad y Estadística

Las observaciones de un estudio específico se enumeran mediante letras minúsculas y subíndices x_i que corresponden a la i -ésima observación de un conjunto $z_1, z_2, \dots, z_i, \dots, z_n$ tal que $z_i < \infty, x_i \in \mathbb{R}$.

Dado un conjunto de observaciones, la variable aleatoria asociada se distingue con la misma letra del alfabeto en mayúscula.

¹⁴ Comisión económica para América Latina y el Caribe, Fondo de las Naciones Unidas para la Infancia, *Pobreza infantil en América Latina y el Caribe*, 2010, p. 165-168.

Una *función de probabilidad (fdp)* será mencionada con las letras f o h , aplicadas a la variable aleatoria Z , tal que la notación resulta:

$$f(z), h(z) \quad (3)$$

Las *funciones de probabilidad acumulada* se asociarán con la misma letra que la función de probabilidad de la cual provienen, sin embargo serán representadas con mayúsculas. Por ejemplo:

Sea $f(z)$ la función de probabilidad asociada a la variable aleatoria Z , entonces $F(z)$ es la función de probabilidad acumulada. La esperanza matemática se presenta con la expresión $E(.) = \mu$, y la varianza con $Var(.) = \sigma^2$.

Para distinguir un *estimador* se utilizará el acento circunflejo ($\hat{\beta}, \hat{\alpha}$) sobre la letra del parámetro a estimar. Puede aplicar para parámetros unidimensionales o vectores de parámetros.

Proporciones

Dado que el objetivo del proyecto es presentar proporciones de la población, es de suma relevancia considerar una notación especial para distinguirlas. Para tal efecto se utilizará la letra π , donde $\pi \in [0,1]$. Esta representación aplica de acuerdo al contexto de la fórmula, tal que si se trata de un vector de probabilidades, entonces

$$\pi = \begin{bmatrix} \pi_1 \\ \vdots \\ \pi_k \end{bmatrix} \text{ donde } \pi_i \in [0,1] \forall i = 1, \dots, k \quad (4)$$

Notación para casos particulares.

En el proyecto se mencionan algunas familias de distribución, las cuales se especifican a continuación:

Cuadro 2.1
Notación para las distribuciones de probabilidad

Familia	Nomenclatura	Parámetros	Notación especial
Normal	$Z \sim N(\mu, \sigma^2)$	$\mu = \text{media}, \sigma^2 = \text{varianza}$	
Normal estándar	$Z \sim N(0,1)$	$\mu = 0; \sigma^2 = 1$	$\Phi(z)$

Familia	Nomenclatura	Parámetros	Notación especial
Bernoulli	$Z \sim Ber(p)$	$p = \text{probabilidad de éxito}$	
Binomial	$Z \sim Bin(n, p)$	$n = \text{intentos totales},$ $p = \text{probabilidad de éxito}$	
Poisson	$Z \sim Poisson(\lambda)$	$\lambda = \text{Número de ocurrencias}$	

Fuente: Elaboración propia

Función liga

La letra g será utilizada exclusivamente para hacer referencia a la función liga de los Modelos Lineales Generalizados. La explicación y detalles de esta función se presentarán posteriormente. Para fines prácticos se referencia a la función liga como *la liga*.

2.3 Planteamiento teórico: Modelos Lineales Generalizados

En la introducción del capítulo 2, se menciona que los Modelos Lineales Simples consideran la relación lineal entre la variable dependiente y la variable independiente siendo el método más recurrido el de mínimos cuadrados.

La ventaja de los Modelos Lineales Simples es la sencillez de la representación de los datos mediante una recta, pues para cualquier valor $Z = z$, la variable Y puede representarse mediante la ecuación $Y = \alpha + \beta z + \varepsilon$, donde α y β son los estimadores de los mínimos cuadrados y ε es una variable aleatoria que mide los errores, tal que $\varepsilon \sim N(0, \sigma^2)$.

2.3.1 Ventajas de los Modelos Lineales Generalizados sobre los Modelos Lineales

De acuerdo a Antoine Guisan¹⁵, los Modelos Lineales Generalizados son una extensión de los Modelos Lineales Simples pues incorporan mayor flexibilidad en el modelo evitando que los datos se sometan a escalas antinaturales, permitiendo la no linealidad con varianza no constante.

Otra de las bondades que presentan los Modelos Lineales Generalizados respecto a la regresión lineal, es la factibilidad de aproximar los datos a distribuciones de la familia exponencial. Abarcan casos especiales como: la regresión lineal, los modelos de análisis de varianzas, los modelos logit y los probit para respuestas de dos categorías, los modelos log-lineales con múltiples respuestas para contar y algunos comúnmente utilizados para datos de supervivencia.

¹⁵ Guisan, Antoine, Edwards, Thomas C., Hastie, Trevor, Generalized linear al generalized additive models in studies of species distributions: setting the scene, Ecological Modelling, Vol. 157, 2002, p. 89-100.

Adicionalmente, los Modelos Lineales Generalizados consideran técnicas para realizar estimaciones más flexibles y precisas mediante los estimadores obtenidos por el método de máxima verosimilitud.

La varianza de una variable Y depende de la esperanza (μ) y mediante la función $v(\mu)$, para todos los casos $Var(Y) = \phi v(\mu)$, con ϕ una escala (conocida como parámetro de dispersión). Entonces si el parámetro es más alto que el correspondiente a la distribución, se dice que los datos se encuentran sobre dispersos. Cuando los datos se encuentran sobre dispersos, se estiman los parámetros con cuasiverosimilitud, una generalización de los mínimos cuadrados; tomando así una ventaja más respecto a los Modelos Lineales Simples.

2.3.2 Componentes de los Modelos Lineales Generalizados

Los Modelos Lineales Generalizados tienen tres componentes: la componente aleatoria, la componente sistemática y la función liga.

Componente aleatoria

La componente aleatoria consiste en una variable aleatoria Z con N observaciones independientes (z_1, z_2, \dots, z_N) tal que Y mide el recuento de las observaciones (z_i) . Asimismo, la variable Y distribuye de acuerdo a la familia exponencial, con esperanza matemática μ .

Componente sistemática

El componente sistemático es también conocido como *predictor lineal*, siendo una función lineal de predictores, tal que

$$\eta_i = \sum_{j=1}^m \beta_j z_{ij} \quad (5)$$

Una variante del modelo considera el uso de variables cualitativas y cuantitativas que producen conjuntos de variables virtuales asociadas a un modelo (0,1) tal que

$$\eta_{ij} = \beta v_{ij} + \gamma u_{ij} \text{ donde } i = 1, \dots, n; j = 1, \dots, p \quad (6)$$

Con lo anterior se describe la posibilidad de modelar *variables creadas* a partir de las variables del modelo para estimar interacciones entre las variables o efectos curvilíneos¹⁶.

Función liga

La función liga o también conocida como *la liga*, es una función invertible que transforma la componente aleatoria a la componente sistemática, tal que

$$\eta_i = \sum_{j=1}^m \beta_j z_{ij} = g(\mu_i) = g(E(Y)) \quad (7)$$

Dado que la función g es invertible, entonces

$$g^{-1}(\eta_i) = \mu_i = g^{-1} \left[\sum_{j=1}^m \beta_j z_{ij} \right] \quad (8)$$

Algunas ligas comúnmente utilizadas se especifican en el cuadro 2.2.

Cuadro 2.2
Algunas funciones liga y sus inversas

Liga	$\eta_i = g(\mu_i)$	$\mu_i = g^{-1}(\eta_i)$
Identidad	μ_i	η_i
Log	$\log_e \mu_i$	e^{η_i}
Inversa	μ_i^{-1}	η_i^{-1}
Inversa cuadrada	μ_i^2	$\eta_i^{-1/2}$
Raíz cuadrada	$\sqrt{\mu_i}$	η_i^2
Logit	$\log_e \frac{\mu_i}{1 - \mu_i}$	$\frac{1}{1 + e^{\eta_i}}$
Probit	$\Phi^{-1}(\mu_i)$	$\Phi(\eta_i)$
Log-log	$-\log_e[-\log_e(\mu_i)]$	$\exp[\exp(-\eta_i)]$
Complemento log-log	$\log_e[-\log_e(1 - \mu_i)]$	$1 - \exp[\exp(\eta_i)]$

Fuente: Applied Regression Analysis and Generalized Linear Models¹⁷.

¹⁶ Nelder J.A. and Wedderburn, W.M., Generalized Linear Models, Journal of the Royal Statistical Society, Series A(General), Wiley for the Royal Society, Vol 135 No. 3, 1972, p. 370-384.

¹⁷ Fox, John, Applied Regression Analysis and Generalized Linear Models, Sage Publications, Segunda Edición 2008, páginas 379-424

Los Modelos Lineales Generalizados se encuentran categorizados de acuerdo a la respuesta a modelar y son las funciones logit, probit, log-log y complemento log-log, mostradas en el cuadro 2.2, las que modelan respuestas binomiales.

2.3.3 Modelos para variables respuesta con dos categorías.

Cuando el componente aleatorio considera respuestas con escala binaria (éxito o fracaso) se tiene a disposición la regresión logística con cuatro modelos: Logit, Probit, log-log o complemento log-log. En el proyecto se utilizarán los modelos Logit y Probit por resultar los más confiables y de mayor uso en la práctica.

Entonces, sea Z_k un conjunto de variables independientes. Sea X una variable aleatoria con respuestas dicotómicas, tal que X es explicada por las variables Z_k y

$$X = \begin{cases} 1 & \text{cuando se trata de un éxito} \\ 0 & \text{si es un fracaso} \end{cases}$$

Donde $P(X = 1) = \pi$ y $P(X = 0) = 1 - \pi$.

Si existen n variables aleatorias X_1, X_2, \dots, X_n tal que $P(X_j = 1) = \pi_j$, entonces la probabilidad conjunta para la variable aleatoria Y , donde $Y = \sum X_j$ es:

$$\prod_{j=1}^n \pi_j^{x_j} (1 - \pi_j)^{1-x_j} = \exp \left[\sum_{j=1}^n x_j \log \left(\frac{\pi_j}{1 - \pi_j} \right) + \sum_{j=1}^n \log (1 - \pi_j) \right] \quad (9)$$

La expresión (9) ratifica que la distribución conjunta de Y pertenece a la familia exponencial, por lo que un Modelo Lineal Generalizado es de utilidad. Para el caso particular donde $\pi_i = \pi_j \forall i \neq j, Y \sim Bin(n, \pi)$.

Para el caso general, donde existen N variables aleatorias independientes Y_1, Y_2, \dots, Y_N donde cada Y_j corresponde al j -ésimo suceso en N subconjuntos de datos, entonces la función de máxima verosimilitud es:

$$L(\pi_1, \pi_2, \dots, \pi_N; y_1, y_2, \dots, y_N) = \left[\sum_{i=1}^N y_i \log \left(\frac{\pi_i}{1 - \pi_i} \right) + n_i \log(1 - \pi_i) + \log \binom{n_i}{y_i} \right] \quad (10)$$

Para los Modelos Lineales Generalizados en los que se estiman proporciones, se asume que $P_i = \frac{Y_i}{n_i}$ para cada subconjunto en términos de factores y otras variables explicativas que caracterizan el subconjunto.

Dado que $E(Y_i) = n_i\pi_i$ entonces, $E(P_i) = \pi_i$, por lo que

$$\eta_i = g(E(P_i)) = g(\pi_i) = z_i^T \beta \quad (11)$$

Donde, z_i es el vector de variables explicativas, β es el vector de parámetros y g es la liga.

Dado que los valores ajustados pueden ser menores a cero y mayores a uno, se utiliza una distribución de tolerancia tal que con la función garantice que los resultados se encontrarán en el intervalo $[0,1]$, siendo modelado esto con una función de distribución acumulada tal que:

$$\pi = \int_{-\infty}^t f(s) ds \quad (12)$$

$$f(s) \geq 0 \text{ y } \int_{-\infty}^{\infty} f(s) = 1 \quad (13)$$

La función $f(s)$ es llamada función de tolerancia. Algunas funciones de tolerancia comúnmente utilizadas son la distribución Uniforme (en el modelo de identidad), la distribución Normal (en el modelo Probit), o el modelo logit que utiliza una función especial presentada en el cuadro 2.2.

Modelo Probit

El *modelo Probit* es un modelo comúnmente utilizado en las ciencias biológicas y sociales. En este modelo, la distribución Normal funge como la distribución tolerancia por lo que

$$\pi = \Phi\left(z - \frac{\mu}{\sigma}\right) \quad (14)$$

Por lo tanto, los predictores lineales para el caso de dos variables explicativas, están dados por

$$\beta_1 = -\frac{\mu}{\sigma}, \beta_2 = \frac{1}{\sigma} \quad (15)$$

Y la liga es la inversa de la distribución Normal estándar aplicada al vector de probabilidades π_i , aplicable con un modelo con n variables independientes:

$$\eta_i = g(E(P_i)) = g(\pi_i) = \Phi^{-1}(\pi) \quad (16)$$

Modelo Logit

El *modelo Logit* o *Logistic* también es ampliamente recurrido y de cálculo más simple. La función logistica está dada por:

$$f(s) = \frac{1}{[1 + \exp(-\beta_0 + \sum \beta_i z_i)]} \quad (17)$$

Tal que

$$\pi = P(x = 1 | z_1, \dots, z_k) = \log_e \frac{1}{1 + \exp(-\beta_0 + \sum \beta_i z_i)} \quad (18)$$

Y la liga

$$\eta_i = g(E(P_i)) = g(\pi_i) = \log\left(\frac{\pi_i}{1 - \pi_i}\right) \quad (19)$$

Los modelos Logit y Probit son los más utilizados en la práctica pues se encuentran implementados en la mayoría de los programas estadísticos, sin embargo existen otros métodos para respuestas dicotómicas explicados por distintos autores.

Dado que el predictor lineal proporciona resultados $P(x = 1|z_1, \dots, z_k)$ y el valor buscado es una probabilidad no condicional, entonces por el Teorema de la Probabilidad Total¹⁸

$$P(X) = \sum P(X|z_1, \dots, z_k)P(z_1, \dots, z_k) \quad (20)$$

2.3.4 Verificación de la adecuación del modelo

Si bien es cierto que el análisis de varianza en los Modelos Lineales Simples resulta de utilidad, para los Modelos Lineales Generalizados resulta conveniente el uso de la devianza o *log likelihood ratio statistic*. Entonces, la devianza se define por¹⁹

$$D = 2(l(b_{max}; y) - l(b; y)) \quad (21)$$

Donde $l(b_{max}; y)$ denota el modelo completo (cuando se utilizan todas las variables con su valor máximo) y $l(b; y)$ cuando se utiliza el modelo con los valores nulos en todas las variables (modelo nulo). Por lo tanto, valores grandes en la devianza indican una pobre descripción en los datos.

La devianza no involucra ningún parámetro por lo que una hipótesis puede probarse utilizando directamente la aproximación

$$D \sim \chi^2(N - p) \quad (22)$$

Donde p es el número de parámetros estimados y N el número de patrones covariados. Este estimador resulta conveniente con estudios de buen tamaño, aunque algunos paquetes estadísticos poseen la robustez para tratar estudios con muestras relativamente pequeñas²⁰.

¹⁸ Rincón, Luis, Introducción a la probabilidad, Facultad de Ciencias, agosto 2013, página 65

¹⁹ Dobson, Annete J., An introduction to generalized linear models, Chapman and Hall/CRC, Segunda Edición, 2001, capítulo 5.

²⁰ Ídem.

El p-value en una de sus interpretaciones es la probabilidad de obtener un resultado al menos tan extremo como el que realmente se ha obtenido²¹. En los análisis de bondad de ajuste a realizar, la prueba para aceptar una hipótesis considera un p-value de 0.05.

Otro estadístico útil para la selección o descarte de un modelo el *Akaike Information Criterion* (AIC). Su cómputo es sencillo pues se define como:

$$AIC = 2k - 2\ln(L) \quad (23)$$

Donde k es el número de parámetros del modelo y L es el valor máximo de la función. El estadístico es una buena medida de un modelo sin utilizarlos datos²². Akaike demostró que si los estimadores máximo verosímiles convergen dado que la muestra crece, entonces, el estimador se aproximará al valor real y el AIC tiende a cero, por lo que en consecuencia el modelo es aceptable.

El valor del AIC es irrelevante como dato único, sin embargo al comparar los resultados de dos modelos aplicados, es sencilla la elección del mejor modelo al verificar el AIC menor.

Convergencia del modelo

Los modelos lineales generalizados utilizan estimadores máximo verosímiles. Dado que es necesario encontrar el valor máximo de la ecuación resultante con n variables y n incógnitas, el método numérico utilizado es el conocido como Newton-Rapson²³.

Como en todos los algoritmos, existe un criterio de paro tal que el número de iteraciones se limita para obtener eficiencia computacional. Para algunos programas estadísticos, entre ellos R y S, el criterio de paro es definido por

$$|\text{devianza}^{(i)} - \text{devianza}^{(i-1)}| < \epsilon(\text{devianza}^{(i-1)} + \epsilon) \quad (24)$$

²¹ Schervish MJ, *P Values: What They Are and What They Are Not*. The American Statistician, Num 50, vol.3,1996,p. 203-206

²² P.Burnham, Kenneth and Anderson,David R ,*Multimodel inference*, Sociological Methods & Research, Vol. 33, Núm. 2, noviembre 2004

²³ Nelder J.A. and Wedderburn, W.M., *Generalized Linear Models*, Journal of the Royal Statistical Society, Series A(General), Wiley for the Royal Society, Vol 135 No. 3, 1972, p. 376

donde el valor por defecto de ϵ es 10^{-424} , de modo tal que si el criterio no se cumple, después de una cantidad determinada de iteraciones, se dice que la aproximación no converge y el modelo no lo hace.

Este tipo de problemas de convergencia, se deben a la inexistencia de estimadores o bien a que el valor máximo se alcanza en dos puntos.

²⁴ Venables, W.N. and Ripley, B.D., *Modern Applied Statistics with S*, Springer, Cuarta edición, 2002, p.188

Capítulo 3

Modelo de estimación de probabilidad de pobreza extrema

Introducción

Después de analizar la concepción de la pobreza, su tratamiento y diversas formas en las que el Estado pretende mitigarla y, tras estudiar los Modelos Lineales Generalizados, sus características, bondades y limitantes, en éste capítulo se conjunta la información presentada en los capítulos previos para concluir con estimaciones de la pobreza extrema.

En los apartados de este capítulo se describe el proceso para demostrar la hipótesis planteada en el Diseño metodológico; asimismo se presentan los resultados obtenidos a nivel nacional y para las entidades federativas identificadas con la mayor proporción de su población en situación de pobreza. Dado que todo modelo requiere ajustes y observaciones también se presentan los resultados del ajuste realizado.

Para finalizar, con el modelo lineal generalizado ajustado se plantea un procedimiento para estimar los indicadores de pobreza extrema nacionales a 2030.

3.1 Propósito y alcance

El propósito del proyecto es comprobar que los indicadores estipulados en la Ley General de Desarrollo Social²⁵ explican la variable *pobreza extrema*, es decir, que la proporción de personas que viven en pobreza extrema depende de los ocho indicadores mínimos estipulados por la Ley. Adicionalmente, en el estudio se incluye la zona de residencia como un factor que influye en la pobreza extrema, pues al analizar las carencias rezago educativo, acceso a servicios médicos y acceso a servicios en la vivienda, resulta coherente considerar que estos servicios son proporcionados por el Estado, y la carencia de alguna de éstas se debe a la lejanía o proximidad de cada vivienda a los municipios o comunidades donde se encuentra el colegio, hospital, clínica o instalación hidráulica más cercana.

²⁵ Ley General de Desarrollo Social, Ciudad de México, Diario Oficial de la Federación, 16 de junio 2012

De acuerdo a los reportes y análisis realizados por CONEVAL²⁶, la situación de una persona se clasifica en más de una categoría: *sin carencias, vulnerable por ingresos, vulnerable por carencias sociales, pobre moderado o pobre extremo*. Dado que el interés de este proyecto es presentar la proporción de personas en el caso de pobreza extrema, es suficiente con agrupar las categorías *sin carencias, vulnerable por ingresos, vulnerable por carencias sociales y pobre moderado* como una categoría *no pobre extremo*, de tal forma que la estimación se realice mediante dos categorías: *pobre extremo o no pobre extremo*.

Sin embargo, por la documentación y la información contenida en las bases de datos, sin fines pretenciosos, es factible, con un esfuerzo adicional, obtener información en una clasificación más amplia, haciendo uso de Modelos Lineales Generalizados para respuestas con más de dos categorías.

3.2 Diseño metodológico para determinar índices de pobreza extrema

Dado que el objetivo de esta tesis es estimar la pobreza extrema, a partir de la definición de la Ley General de Desarrollo Social y de la definición construida por el CONEVAL, con el propósito de construir un modelo lineal generalizado, se detallará en este apartado los pasos a seguir para cumplir el propósito.

Adicionalmente, analizar la pobreza mediante un Modelo Lineal Generalizado permite estimar en distintos momentos la información para proporcionar proyecciones de utilidad para la toma de decisiones en el mediano y largo plazo en materia de políticas gubernamentales, las cuales se enfocan en la atención de la pobreza a nivel nacional.

Definición de variables

De acuerdo a lo estipulado en el artículo 36 de la Ley General de Desarrollo Social, se debe considerar ocho indicadores mínimos para la medición de la pobreza: ingreso per cápita, rezago educativo, acceso a los servicios médicos, seguridad social, alimentación, hacinamiento, servicios en la vivienda y cohesión social. Adicionalmente se considera el indicador de ubicación de la residencia. Estos indicadores corresponden a la componente sistemática.

La componente aleatoria es la proporción asociada a la pobreza extrema. Dado que estamos suponiendo una proporción con dos respuestas (pobre extremo o no pobre extremo), la liga

²⁶ Consejo Nacional de Evaluación de la Política de Desarrollo social, *Lineamientos y criterios generales para la definición, identificación y medición de la pobreza*, Ciudad de México, Diario Oficial de la Federación, 2010

a utilizar será la Logit o la función Probit y los predictores lineales serán las combinaciones de carencias definidas de acuerdo al apartado *Fases de cálculo*.

Hipótesis y supuestos

La hipótesis principal es que la proporción de la población con pobreza extrema es explicada por los ocho indicadores establecidos en la Ley General de Desarrollo Social²⁷ y el indicador añadido para la medición respecto a la zona de residencia.

Por lo tanto las respuestas que deberemos resolver con este proyecto son las siguientes:

- a) Si se tiene rezago educativo, ¿la proporción de la pobreza extrema se afecta?
- b) Si la carencia de servicios médicos se agudiza ¿la proporción de la pobreza extrema se modifica?
- c) La carencia de la seguridad social, ¿afecta la proporción de la población que se encuentra en pobreza extrema?
- d) ¿Cuál es la influencia de la variable *alimentación*, en la proporción de la pobreza extrema?
- e) ¿Cuál es la significancia de la vivienda o los servicios de la vivienda en la proporción de la población en pobreza extrema?
- f) ¿Acaso la cohesión social tiene relevancia en el porcentaje de personas en pobreza extrema?

Tratamiento de los datos

Dada la cantidad de información a recopilar, no es factible considerar el uso de encuestas, entrevistas u observaciones del entorno para el desarrollo del proyecto de tesis. Toda la información será obtenida de encuestas, conteos o censos previamente recopilados por el Instituto Nacional de Geografía y Estadística (INEGI), por lo que considerar el momento y lugar del levantamiento es relevante.

El procedimiento ejecutado para probar la hipótesis son las siguientes:

- Determinación de la fuente de información: Encontrar la información de entre las distintas entidades generadoras. Se seleccionan las bases que guardan características similares para realizar estimaciones sobre información consistente.

²⁷ Ley General de Desarrollo Social, Ciudad de México, Diario Oficial de la Federación, 16 de junio 2012.

- Recopilación de la información: Las bases de datos a disposición del público, y generadas por las distintas entidades gubernamentales, se recopilan y etiquetan para su manipulación.
- Análisis de la información: Se analizan las características de las bases de datos. Este análisis incluye la determinación de campos utilizables, compatibles e información extraíble por base de datos.
- Homologación de la información: Implica la transformación de los campos para crear una base de datos homogénea que represente lo mismo para toda la información.
- Creación de tablas para modelar: Manipular la información para su correcto procesamiento en el programa estadístico.
- Aplicación del modelo: Procesamiento de la información final y tratamiento en el programa estadístico.
- Ajuste del modelo: Validar los resultados y realizar ajustes para reprocesar en caso de ser necesario y aplicar el nuevo modelo para encontrar el estimador idóneo.

Pruebas de factibilidad

En la aplicación de todo modelo es necesario validar que la información aproximada es congruente con la información observada, por lo que es necesario verificar si el ajuste es confiable. Para tal fin, en el caso de los Modelos Lineales Generalizados se utilizan dos herramientas: a) La prueba de bondad de ajuste Chi-cuadrado, b) la Devianza.

El Akaike Information Criterion (AIC) también es una medida de factibilidad aunque no es estadístico de prueba contundente para aceptar o descartar la prueba de hipótesis, sin embargo se utilizará como criterio adicional para la selección del modelo idóneo.

3.3 Diseño, planeación y ejecución

En este apartado se presentan las acciones efectuadas desde la concepción del problema a la determinación del modelo adecuado para comprobar o refutar la hipótesis planteada. Del mismo modo se plantean las actividades a ejecutar para obtener las estimaciones que sirven de fundamento para comprobar la hipótesis.

La planeación considera la determinación de las herramientas computacionales a utilizar para el procesamiento de la información. Las bases de datos se encuentran en formato DBF lo que facilita la manipulación de las tablas y su transformación a los indicadores necesarios. Para estimar el Modelo Lineal Generalizado se utiliza el software estadístico R, por ser de uso libre y por contar con documentación amplia y de fácil acceso. La librería empleada fue *glm* con la que se estiman distintos Modelos Lineales Generalizados. En el apartado 3.4 se construye un

programa en VBA Microsoft Office Excel que calcula 10 000 escenarios, donde el conjunto de observaciones para cada escenario es seleccionado de manera aleatoria de los años estimados en el proyecto (2004, 2006, 2008, 2010 y 2012). Posteriormente se grafican los resultados con la herramienta *Histograma* del mismo Excel.

Definición de la estimación

En una primera fase, se investiga la teoría, los fundamentos y los procesos necesarios para estimar la proporción de la población en situación de pobreza extrema utilizando los Modelos Lineales Generalizados.

Dado que la población se segmenta en dos categorías: persona en situación de pobreza extrema o persona que no se encuentra en situación de pobreza extrema, entonces los Modelos Lineales Generalizados recomendados son el modelo Probit o el modelo Logit, según el ajuste de los datos. Por lo tanto, la pauta para la selección de los modelos se alinea con el objetivo del proyecto. Con la premisa anterior, se construyen las bases de datos que alimentan al paquete estadístico para la modelación.

Entonces, al asociar el Modelo Lineal Generalizado para variables con una respuesta binaria con el problema planteado, se tiene π_j la variable aleatoria asociada a la proporción de pobreza extrema para el año j , donde $j = 2004, 2006, 2008, 2010, 2012$. Sea Z_i la variable aleatoria que mide la carencia del indicador i , tal que Z_i asume alguno de los valores presentados en el cuadro 3.1.

Cuadro 3.1
Definición de valores por indicador de pobreza

Indicador	Definición de valores
Acceso a la alimentación	$Z_1 = \begin{cases} 0 & \text{si no se carece de alimentación} \\ 1 & \text{si se carece de alimentación} \end{cases}$
Calidad de espacios en la vivienda	$Z_2 = \begin{cases} 0 & \text{si no se carece de espacios en la vivienda} \\ 1 & \text{si se carece de espacios en la vivienda} \end{cases}$
Servicios básicos en la vivienda	$Z_3 = \begin{cases} 0 & \text{si no se carece de los servicios básicos en la vivienda} \\ 1 & \text{si se carece de servicios básicos en la vivienda} \end{cases}$
Rezago educativo	$Z_4 = \begin{cases} 0 & \text{si no se padece rezago educativo} \\ 1 & \text{si se padece rezago educativo} \end{cases}$
Acceso a la seguridad social	$Z_5 = \begin{cases} 0 & \text{si carece acceso a la seguridad social} \\ 1 & \text{si no carece de seguridad social} \end{cases}$
Acceso a los servicios médicos	$Z_6 = \begin{cases} 0 & \text{si carece a los servicios médicos} \\ 1 & \text{si no carece de acceso a servicios médicos} \end{cases}$
Ingresos por arriba de la Línea de Bienestar mínimo	$Z_7 = \begin{cases} 0 & \text{si el ingreso es inferior a la línea de bienestar mínima} \\ 1 & \text{si el ingreso es mayor a la línea de bienestar mínima} \end{cases}$

Indicador	Definición de valores
Índice de marginación	$Z_8 = \begin{cases} 0 & \text{si el nivel de marginación es medio, alto o muy alto} \\ 1 & \text{si el nivel de marginación es bajo o muy bajo} \end{cases}$
Zona de residencia	$Z_9 = \begin{cases} 0 & \text{si se habita en zona urbana} \\ 1 & \text{si se habita en zona rural} \end{cases}$
Indicador de pobreza extrema	$X = \begin{cases} 0 & \text{si no es pobre extremo} \\ 1 & \text{si es pobre extremo} \end{cases}$
Estado	e denota la entidad federativa en la que se encuentra el indicador.

Fuente: Elaboración propia

Entonces, sea Y_i la variable aleatoria que mide el número de observaciones de las distintas combinaciones de las variables Z_i , tal que cada $Y_i = \sum Z_i$ y $Y_i \sim \text{Bin}(n_i, \pi_j)$, donde n_i representa el número total de casos por cada observación i y π_j es la probabilidad de éxito de dicha combinación.

Por lo tanto, de acuerdo a la fórmula (9) del capítulo 1,

$$\eta_j = g(E(P_j)) = g(\pi_j) = z_i^T \beta \quad (25)$$

donde el vector de variables explicativas z_i es la combinación de los indicadores y β es el vector de predictores lineales a estimar. Se estima π_j con los modelos Probit y Logit, por lo que las ligas a utilizar son las expresadas en las fórmulas (16) y (19), respectivamente. La selección de los estimadores que mejor ajusten se realiza más adelante.

El modelo a utilizar es:

$$P(X|Z_1 = z_1, \dots, Z_9 = z_9) = \log_e \frac{1}{1 + e^{-(\beta_0 + \sum_{j=1}^n \beta_j z_j)}} \quad (26)$$

Estratificación de la información

En el paso *Determinación de la fuente de información* descrito en las *Fases de Cálculo del Diseño metodológico*, se especifica el origen de la información recopilada para el estudio, la cual, de inicio considero diversas fuentes de acuerdo al indicador a analizar.

Cuadro 3.2
Fuentes de información por indicador

Indicador	Fuente de información
Ingreso per cápita	Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH)
Rezago Educativo	Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) Encuestas en hogares – módulo de capacitación y empleo Censo de Población y vivienda

Indicador	Fuente de información
Accesos a servicios de salud	Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH)
Acceso a la seguridad social	Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) Encuestas en hogares – módulo de Seguridad Social Encuesta Nacional de Ocupación y Empleo (ENOE) Encuesta Nacional de empleo
Alimentación	Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH)
Vivienda – Calidad en los espacios, hacinamiento y servicios básicos	Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH)
Cohesión social	Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) Censo de Población y vivienda Índices de Gini publicados por el Banco Mundial

Fuente: Elaboración propia.

Dado que existe una amplia gama de encuestas que proporcionan información, es necesario verificar que los datos corresponden a condiciones similares (lugar y tiempo) para garantizar que la estimación posea un alto grado de confiabilidad.

La información existe en dos formatos: datos consolidados y respuestas a los cuestionarios de las encuestas citadas en el cuadro 3.2. Para los Censos y los Conteos de Población y Vivienda solamente existe información consolidada a disposición del público, por lo que son de poca utilidad para el desarrollo del proyecto.

Las bases de datos se recopilan, clasifican y etiquetan de acuerdo a: la fuente de información generadora, el año de levantamiento de la encuesta y el indicador al que satisface. El cuadro 3.3 resume la disponibilidad de la misma de acuerdo a cada una de las etiquetas.

Cuadro 3.3
Origen de la información disponible

Indicador	Origen
✓	Encuesta Nacional de Ingresos y Gastos en Hogares del INEGI
✓	Censos o Conteos Nacionales, INEGI
✓	Módulo de Capacitación de la Encuesta Nacional de Ingresos y Gastos en Hogares, INEGI.
✓	Módulo de Seguridad Social de la Encuesta Nacional de Ingresos y Gastos en Hogares, INEGI.
✓	Módulo de Seguridad Social de la Encuesta Nacional de Ingresos y Gastos en Hogares, INEGI.

Fuente: Elaboración propia

En el cuadro 3.4 destaca la disponibilidad de información de la Encuesta Nacional de Ingresos y Gastos de los Hogares siendo la encuesta que satisface la mayor cantidad de años y variables para el estudio, mientras que en las encuestas adicionales las muestras tienen diferencias en el levantamiento tanto en tiempo y como en espacio, para el mismo indicador.

Cuadro 3.4
Información disponible por año de levantamiento.

Indicador	Año calendario																	
	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2012
Ingreso per capita	✓		✓		✓		✓		✓		✓	✓	✓		✓		✓	✓
Rezago educativo	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓		✓	✓	✓	✓
Acceso a los servicios de salud			✓				✓				✓		✓		✓	✓	✓	✓
Seguridad Social			✓				✓				✓		✓		✓	✓	✓	✓
Calidad y espacios de vivienda	✓		✓		✓		✓		✓		✓	✓	✓		✓		✓	✓
Servicios básicos vivienda	✓		✓		✓		✓		✓		✓		✓		✓		✓	✓
Alimentación															✓		✓	✓
Grado de cohesión															✓		✓	✓

Fuente: Elaboración propia.

Al analizar la información de las distintas fuentes (encuestas) o la misma pero con levantamientos de distintos años se detectaron diferencias, las cuales se agrupan en tres rubros:

- a. Diferencias en la estructura de la base de datos
 Para una encuesta con levantamientos en años distintos, la información recolectada es la misma pero la composición o por tabla creada está en presentación o formato distinto. Estos cambios se anunciaron y han sido parte del trabajo del INEGI al impulsar la normalización de las tablas y llevarlas a estructuras acordes a las bases de datos relacionales, facilitando así su manipulación, comprensión y publicación.
- b. Diferencias en la información contenida en los campos de la base de datos para levantamientos en años diferentes.
 Pese a que los campos tienen el mismo nombre y la respuesta a la que responden es la misma, no contienen el mismo tipo de dato por un cambio en el valor. Se detectan los siguientes cambios entre ciertos campos de las bases de datos:
 - a. Catálogos diferentes (ampliados o recortados) con mismo tipo de datos.
 - b. Campos que cambiaron su tipo de dato (números a catálogos o viceversa).
- c. Diferencias en el número de campos de las tablas.
 Con la necesidad de recolectar información significativa y relevante, los cuestionarios aplicados tuvieron cambios en los años, provocando modificaciones en algunas tablas.

Estas modificaciones se resumen a la inclusión o eliminación de ciertos campos, dificultando la unión o empate de las tablas de años distintos.

Adicionalmente, las bases de datos de diferentes encuestas son:

- a. Distintas en estructura o características de campos similares,
- b. Consideran información, que por el levantamiento mismo, no es sencillo empatar para tener una muestra homogénea
- c. De toda la información contenida, solo resulta útil un campo o dos que no generan diferencia en el indicador a estimar, pues la información ya se encuentra contenida en alguna encuesta más amplia para años distintos o inclusive, para el levantamiento del mismo año.

Por lo anterior, tras recopilar la información y etiquetarla se establece a la Encuesta Nacional de Ingresos y Gastos de los Hogares para los años 2004, 2006, 2008, 2010 y 2012, como la fuente de información a utilizar pues contiene la información para la mayoría de los indicadores, así como la homogeneidad para su tratamiento.

La ENIGH contiene información a nivel persona – hogar – vivienda. La premisa de la que parte la encuesta es considerar que ciertos parámetros corresponden a la persona (nivel educativo, acceso a servicios médicos, acceso a la seguridad social), otros al hogar (acceso a la alimentación, uso de enseres domésticos) y finalmente algunos sólo son aplicables a la vivienda (zona de residencia, características y servicios básicos de la vivienda), por lo que las estimaciones consideradas en el proyecto corresponden al nivel persona heredando de hogares o vivienda la información correspondiente, para así estimar la proporción de personas en pobreza extrema.

Dada la complejidad de cada base de datos, se transforman las respuestas de los cuestionarios a indicadores de acuerdo a los criterios establecidos en los *Lineamientos y criterios generales para la definición, identificación y medición de la pobreza*²⁸.

Por ingreso per cápita

Se categoriza de acuerdo a la estimación del monto de la canasta básica publicada por CONEVAL a mayo del 2013 (1 179.30 pesos). Para hacer comparable los años (2004, 2006, 2008, 2010) fue necesaria la actualización del ingreso.

²⁸ Consejo Nacional de Evaluación de la Política de Desarrollo social, *Lineamientos y criterios generales para la definición, identificación y medición de la pobreza*, Ciudad de México, Diario Oficial de la Federación, 2010.

El índice más recurrido para realizar las actualizaciones es el Índice Nacional de Precios y Cotizaciones (INPC), sin embargo al índice se encuentra asociado un factor de estacionalidad el cual desvirtúa ligeramente el ingreso capturado en la encuesta. Por lo tanto, se utilizó el Indicador Global de la Actividad Económica (IGAE) pues para la elaboración del Indicador se utiliza el esquema conceptual y metodológico de la contabilidad nacional, mismo que sigue el cálculo del Producto Interno Bruto (PIB) trimestral, sin embargo no incluye a todas las actividades económicas y por ello debe considerársele como un indicador de la tendencia o dirección de la actividad económica en el país para el corto plazo.

Adicionalmente, es importante destacar que el ingreso declarado en las encuestas es un dato proporcionado por el encuestado y se desvirtúa por la desconfianza que una pregunta de índole confidencial es efectuada por un desconocido (encuestador) por lo que, exclusivamente se calcula considerando lo contenido en la tabla, sin corrección de datos.

Por la calidad y espacios de vivienda

Se aplica el criterio definido en los lineamientos del Diario Oficial de la Federación²⁹ establecidos en el Capítulo III, artículo décimo cuarto, en materia de calidad y espacios de la vivienda. La población que resida en viviendas con al menos una de las siguientes características se considera como carente:

- El material de la mayor parte de los pisos es tierra.
- El material de la mayor parte del techo es lámina de cartón o desechos.
- El material de la mayor parte de los muros es embarro o bajareque; carrizo, bambú o palma; lámina de cartón, metálica o de asbesto; o material de desecho.
- Con hacinamiento.

Los materiales de techos y muros se discriminan de acuerdo a lo contenido en las encuestas. El hacinamiento se calcula como el cociente entre el número de residentes de la casa y el número habitaciones de la vivienda. Se considera una carencia si el cociente resultante es mayor a 2.5.

²⁹ Consejo Nacional de Evaluación de la Política de Desarrollo social, *Lineamientos y criterios generales para la definición, identificación y medición de la pobreza*, Ciudad de México, Diario Oficial de la Federación, 2010.

Por acceso a servicios básicos de la vivienda

Los lineamientos³⁰ para la medición de la vivienda consideran cuatro servicios básicos para considerar que una vivienda no es pobre: contar con energía eléctrica, el agua se obtiene de alguna fuente con flujo regular o fácil acceso, drenaje y combustible para cocinar. La carencia de uno de estos servicios implica carencia para este indicador.

Por rezago educativo

La medición del rezago educativo se realiza de acuerdo a lo establecido con los lineamientos³¹ (edad y nivel educativo concluido), sin embargo calcula de acuerdo a la reforma en el artículo 3 de la Constitución de los Estados Unidos Mexicanos³², donde se establece que todos los habitantes del país tienen derecho a la educación media superior.

Al realizar la modificación se limita la comparación respecto los años previos a la última reforma y en consecuencia los indicadores varían respecto a las mediciones oficiales realizadas por el CONEVAL para años previos a la elaboración de este trabajo.

Por servicios médicos

La evaluación considera la accesibilidad y cercanía a servicios médicos ya sean particulares o prestados por paraestatales, independientemente de si hacen o no uso de ellos. También se considera como criterio para la no carencia, el acceso mediante prestadores de servicios asociados a la contratación de una póliza de Gastos Médicos Mayores, ya sea por adquisición propia o como parte de un paquete de prestaciones proporcionadas por la empresa en la que se labora.

Las ENIGH previas a 2008 no contienen esta información, sin embargo el indicador se induce por medio de las preguntas de prestaciones en el trabajo, aunque el indicador es sesgado pues no todas las personas conocen sus prestaciones y algunas personas cuentan con paquetes de servicios médicos proporcionados por particulares (empleados de la Banca de segundo nivel, paraestatales, etc.).

³⁰ Ídem

³¹ Consejo Nacional de Evaluación de la Política de Desarrollo social, *Lineamientos y criterios generales para la definición, identificación y medición de la pobreza*, Ciudad de México, Diario Oficial de la Federación, 2010.

³² *Constitución Política de los Estados Unidos Mexicanos*, Ciudad de México, Diario Oficial de la Federación, 5 de febrero de 1917, Última reforma publicada 26 de febrero, 2013

Por acceso a la alimentación

El cuestionario asociado al rubro de alimentación, considera 16 preguntas a nivel hogar donde se plasman las percepciones de la persona entrevistada. En los lineamientos³³ la escala para la medición de la carencia de alimentación está dada por el grado de carencia, considerando el indicador como una carencia cuando la falta de alimentación se estima severa o grave. Sin embargo, el criterio aplicado en el proyecto considera que si alguna de las preguntas es contestada afirmativamente, entonces la persona es carente de acceso a la alimentación.

La razón por la que se considera como carente de alimentación para cualquier respuesta afirmativa se debe a: a) la ambigüedad de las preguntas, pues miden la percepción; b) las costumbres alimentarias de un hogar a otro varían y c) existe una gran discusión respecto al que debe considerarse como *alimentación mínima* y su escala de medida (alimentos o aportación calórica de los alimentos a los que se tiene acceso en la población en la que se habita).

Por acceso a la seguridad social

Se entiende por acceso a la seguridad social, la conformación de un fondo o ahorro para el retiro, si se trata de una persona activa laboralmente, o bien, la recepción de algún apoyo económico por parte del gobierno o una pensión, para las personas que alcanzaron la cesantía o vejez.

Esta variable se encuentra vinculada con el acceso a servicios médicos, pues las instituciones de seguridad social son las mismas que prestan el servicio médico (IMSS, ISSSTE, ISSSFAM, entre otras). Adicionalmente, se considera como cubierto el acceso si se contrata de manera voluntaria una póliza con componente de ahorro para el retiro.

Para 2004 y 2006 se obtiene esta información con la respuesta de si el encuestado posee o no la prestación del servicio médico proporcionado por alguna paraestatal, pues en el cuestionario no existían preguntas respecto a beneficios del gobierno o contratación particular de un plan de pensiones.

Por el grado de cohesión social

Para establecer el grado de cohesión social el CONEVAL considera las redes sociales y la información proporcionada por el Consejo Nacional de Población (CONAPO) de acuerdo a: el

³³ Op.Cit. Consejo Nacional de Evaluación de la Política de Desarrollo Social

índice de Gini, el grado de polarización social de la entidad federativa o del municipio, la razón de ingreso de la población pobre multidimensional extrema respecto a la población no pobre multidimensional y no vulnerable y el índice de percepción de redes sociales.

Dado que replicar el índice citado por la Ley³⁴ es complejo y desconocido, se utiliza el indicador de marginación calculado por el INEGI en cada ENIGH. Esta información se transforma considerando que un nivel de marginación medio, alto y muy alto implican una carencia social o, en términos de personas, se trata de individuos marginados.

Por la zona de residencia

En la información contenida en los códigos postales del Servicio Postal Mexicano (SEPOMEX) es fácilmente identificable si una colonia o asentamiento se encuentra en una zona rural o urbana. Sin embargo, el levantamiento de la información a nivel vivienda en la ENIGH no considera tal detalle pese a que se incluye la ubicación geográfica. Para mitigar tal situación se consideró como indicador al tamaño de población como tal.

Es así como el número de habitantes por localidad proporciona una referencia respecto a la ubicación de residencia y se asumió como criterio de zona rural a las poblaciones con menos de 15 000 habitantes.

3.3.1 Ejecución del modelo

Inicialmente fue necesario establecer por persona aquella que se encuentra en situación de pobreza extrema. Para identificar a una persona en tal situación se utiliza la definición planteada en los lineamientos establecidos por CONEVAL³⁵: el número de carencias debe ser mayor o igual a tres y el ingreso promedio por persona debe ser inferior a la línea de pobreza. La zona de residencia incluida en el estudio no se considera como una carencia social y como tal, no se contabiliza para determinar pobreza extrema por este rubro, sin embargo se mantiene para evaluar la aportación que tiene al modelo.

La información se encuentra a nivel cuestionario por lo que se asigna a cada registro el valor de X , tal como definido en el cuadro 3.1. Entonces, la variable aleatoria Y contabiliza el número de personas por situación (pobre extremo: $X = 0$, no pobre extremo: $X = 1$) de

³⁴ *Ley General de Desarrollo Social*, Ciudad de México, Diario Oficial de la Federación, 16 de junio 2012.

³⁵ Consejo Nacional de Evaluación de la Política de Desarrollo social, *Lineamientos y criterios generales para la definición, identificación y medición de la pobreza*, Ciudad de México, Diario Oficial de la Federación, 2010

acuerdo a la combinación de carencias ($Z_k, k = 1, \dots, 9$), agrupadas por entidad federativa y año de levantamiento de la información.

El modelo aplicado considera a Y_i como una variable aleatoria Binomial por lo que

$$Y_i = \sum x_j \quad (27)$$

Entonces, al aplicar el Modelo Lineal Generalizado el resultado obtenido es

$$P(X|Z_2, \dots, Z_9) = \frac{1}{1 + e^{-(\beta_0 + \sum_{j=2}^9 \beta_j z_{ij})}} \quad (28)$$

donde la interpretación del resultado es *la proporción de personas en pobreza extrema dada la combinación de carencias respecto a la combinación misma*. La ejecución del modelo proporciona estas probabilidades a manipular posteriormente para obtener el índice de pobreza extrema.

Al ejecutar el modelo para 2004, no se contempla la variable Z_1 (alimentación) dado que no existe con información al respecto, tal como mostrado en el cuadro 3.4. En consecuencia, la ejecución de los modelos Probit y Logit asociados no convergen, y la significancia de todas las variables no satisface el p-value menor a 0.05, visible en el cuadro 3.5. Adicionalmente, la variable Z_6 (servicios médicos) no es estimada por singularidades, consecuencia de la igualación con la variable asociada al acceso a la seguridad social, por la carencia de la información en las encuestas para los años 2004 y 2006.

Cuadro 3.5
Resultados de la aplicación del MLG con 7 variables - 2004

Variables	Logit			Probit		
	Coficiente	Error estándar	p-value	Coficiente	Error estándar	p-value
β_1	-195.2	20860	0.993	-55.57	3608	0.988
e	0.02739	239.6	1	0.004159	41.47	1
Z_2	43.46	7597	0.995	12.37	1316	0.993
Z_3	42.34	7052	0.995	12.16	1181	0.992
Z_4	42.93	6088	0.994	12.29	1079	0.991
Z_5	-84.01	13020	0.995	-24.27	2348	0.992
Z_6	NA	NA	NA	NA	NA	NA
Z_7	173.4	16870	0.992	49.37	2916	0.986

Variables	Logit			Probit		
	Coeficiente	Error estándar	p-value	Coeficiente	Error estándar	p-value
Z ₈	43.49	6502	0.995	12.37	1126	0.991
Z ₉	0.3595	4284	1	0.06179	744.2	1

Fuente: elaboración propia

En el análisis de varianza para el año 2004 destaca la buena aproximación que se realiza para el modelo exceptuando la variable Z₉ asociada a la zona de residencia. De lo descrito en los cuadros 3.5 y 3.6 se concluye que es necesario ajustar el modelo a un menor número de variables.

Cuadro 3.6
Resultados del análisis de bondad de ajuste al MLG con 7 variables - 2004

Variable	Modelo Logit			Modelo Probit		
	Grados de libertad	Deviance	p-value	Grados de libertad	Deviance	p-value
e	1	0.16	0.6881	1	0.16	0.6871
Z ₂	1	69.09	< 2.2e-16	1	69.12	< 2.2e-16
Z ₃	1	43.04	5.4E-11	1	43.04	5.4E-11
Z ₄	1	63.05	2E-15	1	63.09	2E-15
Z ₅	1	26.44	2.7E-07	1	26.37	2.8E-07
Z ₆	0	0		0	0	
Z ₇	1	2108.05	< 2.2e-16	1	2108.05	< 2.2e-16
Z ₈	1	143.59	< 2.2e-16	1	143.59	< 2.2e-16
Z ₉	1	0.27	0.6002	1	0.27	0.6005

Fuente: Elaboración propia

Para ajustar el modelo de 2004 se realizan corridas adicionales retirando las variables Z₅ (acceso a la seguridad social) y Z₆ (acceso a los servicios médicos) en dos escenarios donde la convergencia del modelo se obtuvo al eliminar ambas variables. Estas variables se retiraron porque no existe en la encuesta una pregunta directa para obtener la información, por lo que los supuestos bajo los que se obtuvo el indicador desvirtúan la medición. Por lo tanto, se mantiene el modelo con cinco variables presentado en el cuadro 3.7 – Corrida 2. Eliminar más variables del modelo se contrapone con las premisas de la metodología y se conservan con fines meramente explicativos, sin embargo, al revisar el p-value de las variables se comprueba la hipótesis donde no todas las variables definidas en la Ley General de Desarrollo Social³⁶ aportan a la medición de la pobreza extrema.

³⁶ Ley General de Desarrollo Social, Ciudad de México, Diario Oficial de la Federación, 16 de junio 2012.

Cuadro 3.7
Resultados de la aplicación del MLG eliminando variables - 2004

Indicador	Corrida 1				Corrida 2			
	Modelo Logit		Modelo Probit		Modelo Logit		Modelo Probit	
	Coeficiente	p-value	Coeficiente	p-value	Coeficiente	p-value	Coeficiente	p-value
β_1	-195.2	0.993	-55.57	0.988	-86.26799	0.99	-28.84	0.98
e	0.02739	1	0.004159	1	-0.03595	0.425	0.004982	0.779
Z_2	43.46	0.995	12.37	0.993	25.44297	0.995	7.984	0.991
Z_3	42.34	0.995	12.16	0.992	24.18928	0.995	7.652	0.991
Z_4	42.93	0.994	12.29	0.991	7.91612	1.83E-07	3.39	7.01E-13
Z_5	-84.01	0.995	-24.27	0.992				
Z_6								
Z_7	173.4	0.992	49.37	0.986	82.36431	0.99	26.75	0.982
Z_8	43.49	0.995	12.37	0.991	6.67892	4.99E-08	3.082	6.08E-14
Z_9	0.3595	1	0.06179	1	0.98219	0.256	0.5211	0.118

■ Variable sin considerar en la corrida

Fuente: Elaboración propia

Análogamente, para el 2006 se aplica el modelo sin considerar la variable Z_1 que mide la carencia de alimentación, pues no existe información. Por lo tanto, el modelo a utilizar es:

$$Y_i = \sum x_j \quad (29)$$

Del mismo modo que en la estimación inicial para 2004, el modelo a utilizar ya incluye la variable de alimentación por lo tanto,

$$P(X|Z_1 = z_1, \dots, Z_9 = z_9) = \frac{1}{1 + e^{-(\beta_0 + \sum_{j=1}^n \beta_j z_{ij})}} \quad (30)$$

El modelo diverge por la poca significancia de las variables, tal como se presenta en el cuadro 3.8, por lo que se ejecutan corridas adicionales eliminando las mismas variables que para la información de 2004: acceso a los servicios médicos (Z_6) y acceso a la seguridad social (Z_5), mostrados en el cuadro 3.9.

Cuadro 3.8
Resultados de la aplicación del MLG con 8 variables - 2006

Indicador	Modelo logit			Modelo probit		
	Coeficiente	Error estandar	p-value	Coeficiente	Error estándar	p-value
β_1	-195.9	19540	0.992	-55.69	3388	0.987
e	0.006319	224.6	1	0.001114	38.87	1
Z_2	43.47	7222	0.995	12.37	1244	0.992
Z_3	43.04	6825	0.995	12.29	1162	0.992
Z_4	43.31	6321	0.995	12.34	1093	0.991
Z_5	-86.96	11950	0.994	-24.75	2104	0.991
Z_6	NA	NA	NA	NA	NA	NA
Z_7	174.4	16480	0.992	49.54	2844	0.986
Z_8	43.56	6361	0.995	12.38	1103	0.991
Z_9	0.2318	4148	1	0.04199	716.4	1

Fuente: Elaboración propia

Cuadro 3.9
Resultados de la aplicación del MLG eliminando variables - 2006

Variable	Corrida 1				Corrida 2			
	Modelo logit		Modelo probit		Modelo logit		Modelo probit	
	Coeficiente	p-value	Coeficiente	p-value	Coeficiente	p-value	Coeficiente	p-value
β_1	-195.9	0.992	-55.69	0.987	-53.8133	0.9811	-18.73889	0.9642
e	0.006319	1	0.001114	1	0.02349	0.2631	0.01681	0.1241
Z_2	43.47	0.995	12.37	0.992	21.29014	0.9899	6.70083	0.9825
Z_3	43.04	0.995	12.29	0.992	5.11795	0.00000161	2.42565	3.3E-09
Z_4	43.31	0.995	12.34	0.991	3.18636	3.53E-11	1.53416	1.76E-11
Z_5	-86.96	0.994	-24.75	0.991				
Z_6								
Z_7	174.4	0.992	49.54	0.986	51.31676	0.982	17.42833	0.9667
Z_8	43.56	0.995	12.38	0.991	3.05735	2.19E-11	1.46309	3.08E-11
Z_9	0.2318	1	0.04199	1	0.68609	0.0764	0.40793	0.0413

■ Variable sin considerar en la corrida

Fuente: Elaboración propia

En el año 2008, el INEGI reestructuró los cuestionarios y añadió preguntas para recopilar la información necesaria para estimar la pobreza de acuerdo a los lineamientos de la Ley General de Desarrollo Social³⁷. Por lo tanto, el modelo a aplicar para 2008 ya considera el uso de la variable Z_1 , asociada a la medición de la alimentación.

Entonces, al ejecutar el modelo en el programa estadístico se obtuvo una divergencia similar a la presentada con la información de 2004 y 2006, presente en la corrida 1 del cuadro 3.10.

³⁷ Ley General de Desarrollo Social, Ciudad de México, Diario Oficial de la Federación, 16 de junio 2012.

Para encontrar el modelo es necesario retirar variables y ejecutar nuevas corridas. Con el objeto de homologar los modelos y pese a no ser la variable con significancia menor, se retiró la variable Z_1 (acceso a la alimentación). La razón por la que se retiró esta variable y no otras se debe al cuestionario y la redundancia de las preguntas, las cuales obtienen respuestas subjetivas y más que medir la alimentación miden la percepción. Con esta acción, el modelo converge pero las variables Z_2, \dots, Z_8 continúan siendo insignificantes.

Se calculan corridas adicionales combinado el retiro de las mismas variables eliminadas en los modelos para 2004 y 2006. De lo anterior se obtienen resultados a analizar posteriormente para determinar el modelo recomendado. El resumen de las corridas por combinación de variables y modelo se presenta en el cuadro 3.10 y 3.11.

Dado que los cuestionarios, respuestas y tablas finales son similares para 2010 y 2012, se ajusta la información de los años citados con el mismo criterio. Consistentemente, los resultados para las corridas con la totalidad de las variables resulta divergente, por lo que se evalúa retirando las variables asociadas a alimentación, acceso a servicios médicos y acceso a la seguridad social. Los resultados se presentan en los cuadros 3.12, 3.13 para 2010 y 3.14 y 3.15 para 2012.

Al ejecutar los modelos Probit y Logit en las variables resultantes se presenta un modelo convergente donde las variables que no tienen significancia son Z_7 y Z_9 para 2008, 2010 y 2012. Pese a no tener significancia se conservan las variables dado que la metodología establecida por el CONEVAL considera al ingreso como un determinante de la pobreza extrema. En conclusión, el ingreso económico tiene menor relevancia de la esperada para la pobreza extrema.

En todas las corridas ejecutadas se estiman los coeficientes mediante el modelo Logit y el modelo Probit. El propósito de esta acción es evaluar la efectividad de cada modelo para el conjunto de datos proporcionados.

Cuadro 3.10
Aplicación del modelo lineal generalizado Logit - 2008
Coefficientes y p-value resultantes de la ejecución del modelo en varias combinaciones de variables

<	Corrida 1		Corrida 2		Corrida 3		Corrida 4		Corrida 5	
	Coeficiente	p-value								
β_1	-241	0.991	-103.1	0.982	-28.40534	0.958	-27.617716	0.956	-36.201169	0.96227
e	0.007859	1	-0.00616	0.633	0.004195	0.686	-0.001781	0.868	0.001218	0.92204
Z_1	44.34	0.992								
Z_2	43.93	0.997	19.62	0.985	1.187598	0.138	1.084787	0.187	3.397118	0.00197
Z_3	43.64	0.994	22.21	0.983	3.558527	5E-14	3.826574	1E-15	5.406405	5.43E-16
Z_4	44.26	0.992	21	0.984	2.221208	< 2e-16	2.385	< 2e-16	4.301956	< 2e-16
Z_5	-44.4	0.992	-21	0.984					-4.453759	< 2e-16
Z_6	-42.74	0.993	-20.45	0.985			-2.287247	2E-08		
Z_7	218	0.989	104	0.982	26.3048	0.961	27.623755	0.956	33.183817	0.96542
Z_8	44.04	0.993	20.73	0.984	2.217825	< 2e-16	2.42574	< 2e-16	4.077731	5.37E-16
Z_9	-0.02743	1	0.1239	0.621	0.19891	0.311	0.330398	0.104	-0.035714	0.87952

Variable no considerada en la corrida

Fuente: Elaboración propia

Cuadro 3.11
Aplicación del modelo lineal generalizado Probit - 2008
Coefficientes y p-value resultantes de la ejecución del modelo en varias combinaciones de variables

Variables	Corrida 1		Corrida 2		Corrida 3		Corrida 4		Corrida 5	
	Coeficiente	p-value								
β_1	-68.29	0.986	-29.5245	0.954	1.136724	0.914	-11.163644	0.903	-14.62	0.91832
e	0.001465	1	-0.00414	0.607	0.001772	0.766	-0.001451	0.813	0.0009954	0.89295
Z_1	12.51	0.988								
Z_2	12.44	0.995	5.189399	0.965	0.686613	0.139	0.632344	0.187	1.743	0.00215
Z_3	12.39	0.99	6.788876	0.955	1.971601	<2e-16	2.14025	< 2e-16	2.856	< 2e-16
Z_4	12.5	0.988	6.050483	0.959	1.335071	<2e-16	1.4281	< 2e-16	2.212	< 2e-16
Z_5	-12.52	0.988	-6.04961	0.959					-2.304	< 2e-16
Z_6	-12.21	0.989	-5.70401	0.962			-1.373477	3E-09		
Z_7	61.87	0.981	30.0953	0.953	9.912618	0.924	11.201387	0.902	13.14	0.9266
Z_8	12.46	0.989	5.880063	0.961	1.327902	<2e-16	1.45031	< 2e-16	2.099	< 2e-16
Z_9	-0.0075	1	0.073734	0.637	0.105636	0.349	0.176979	0.127	-0.04804	0.73281

Variable no considerada en la corrida

Fuente: Elaboración propia

Cuadro 3.12
Aplicación del modelo lineal generalizado Logit - 2010
Coefficientes y p-value resultantes de la ejecución del modelo en varias combinaciones de variables

Variables	Corrida 1		Corrida 2		Corrida 3		Corrida 4		Corrida 5	
	Coefficiente	p-value	Coefficiente	p-value	Coefficiente	p-value	Coefficiente	p-value	Coefficiente	p-value
β_1	-237.5	0.985	-102.3	0.9746	-29.627273	0.938	-29.942126	0.959	-36.71302	0.9466
e	-0.008763	1	-0.009081	0.4119	-0.00623	0.476	-0.001885	0.836	-0.01363	0.2009
Z_1	43.68	0.989								
Z_2	43.45	0.989	20.85	0.9775	2.771666	<2e-16	3.013883	< 2e-16	4.5281	<2e-16
Z_3	43.42	0.991	21.39	0.9769	3.494332	<2e-16	3.604854	< 2e-16	5.22001	<2e-16
Z_4	43.63	0.989	21.04	0.9773	2.817535	<2e-16	3.022099	< 2e-16	4.75504	<2e-16
Z_5	-43.9	0.99	-20.56	0.9778					-4.52024	<2e-16
Z_6	-41.16	0.991	-20.44	0.9779			-3.21164	9.54E-13		
Z_7	213.9	0.984	102.6	0.9745	26.463199	0.945	29.469015	0.959	33.11452	0.9519
Z_8	43.43	0.989	20.66	0.9777	2.410018	<2e-16	2.629517	< 2e-16	4.36593	<2e-16
Z_9	-0.1242	1	-0.3508	0.0913	0.041527	0.801	0.107681	0.528	-0.41411	0.0404

Variable no considerada en la corrida

Fuente: Elaboración propia

Cuadro 3.13
Aplicación del modelo lineal generalizado Probit - 2010
Coefficientes y p-value resultantes de la ejecución del modelo en varias combinaciones de variables

Variables	Corrida 1		Corrida 2		Corrida 3		Corrida 4		Corrida 5	
	Coefficiente	p-value	Coefficiente	p-value	Coefficiente	p-value	Coefficiente	p-value	Coefficiente	p-value
β_1	-67.65	0.975	-31.037569	0.9559	-12.042171	0.866	-12.506131	0.898	-15.284084	0.8775
e	-0.001662	1	-0.005706	0.4028	-0.003746	0.453	-0.001291	0.804	-0.008315	0.1817
Z_1	12.39	0.982								
Z_2	12.35	0.982	6.319226	0.9613	1.577531	<2e-16	1.74075	< 2e-16	2.375814	<2e-16
Z_3	12.35	0.985	6.655622	0.9592	1.985932	<2e-16	2.070511	< 2e-16	2.787471	<2e-16
Z_4	12.38	0.982	6.43938	0.9606	1.613707	<2e-16	1.756899	< 2e-16	2.514885	<2e-16
Z_5	-12.43	0.984	-6.140236	0.9624					-2.425898	<2e-16
Z_6	-11.92	0.984	-6.059677	0.9629			-1.818172	3.55E-13		
Z_7	61.13	0.973	31.177493	0.9557	10.240357	0.886	12.175738	0.901	13.491614	0.8918
Z_8	12.35	0.982	6.202111	0.962	1.369584	<2e-16	1.529198	< 2e-16	2.286703	<2e-16
Z_9	-0.02516	1	-0.219106	0.0877	0.016492	0.861	0.053041	0.586	-0.247397	0.0378

Variable no considerada en la corrida

Fuente: Elaboración propia

Cuadro 3.14
Aplicación del modelo lineal generalizado Logit - 2012
Coefficientes y p-value resultantes de la ejecución del modelo en varias combinaciones de variables

Variables	Corrida 1		Corrida 2		Corrida 3		Corrida 4		Corrida 5	
	Coefficiente	p-value	Coefficiente	p-value	Coefficiente	p-value	Coefficiente	p-value	Coefficiente	p-value
β_1	-235.2	0.987	-101.7	0.978	-32.09357	0.96397	-31.0558	0.96336	-39.302134	0.95
e	0.005172	1	0.006265	0.667	0.01728	0.12429	0.01765	0.12079	0.006656	0.642
Z_1	43.23	0.991								
Z_2	43.21	0.991	20.85	0.98	3.28972	< 2e-16	3.37089	< 2e-16	5.301594	< 2e-16
Z_3	42.48	0.992	20.7	0.98	3.01473	1.79E-14	3.11583	5.13E-15	5.152446	5.2E-13
Z_4	43.15	0.991	20.68	0.98	3.16298	< 2e-16	3.2189	< 2e-16	5.174992	< 2e-16
Z_5	-43.41	0.991	-20.65	0.98					-5.126432	< 2e-16
Z_6	-40.41	0.992	-20.1	0.981			-2.13218	0.00134		
Z_7	211	0.986	101.5	0.978	28.06815	0.96848	29.02819	0.96575	34.599018	0.956
Z_8	43.43	0.991	20.43	0.98	2.71169	< 2e-16	2.71471	< 2e-16	4.985992	1.11E-15
Z_9	-0.05011	1	0.3407	0.209	0.57689	0.00577	0.5855	0.00579	0.310818	0.242

Variable no considerada en la corrida

Fuente: Elaboración propia

Cuadro 3.15
Aplicación del modelo lineal generalizado Probit - 2012
Coefficientes y p-value resultantes de la ejecución del modelo en varias combinaciones de variables

Variables	Corrida 1		Corrida 2		Corrida 3		Corrida 4		Corrida 5	
	Coefficiente	p-value	Coefficiente	p-value	Coefficiente	p-value	Coefficiente	p-value	Coefficiente	p-value
β_1	-67.23	0.979	-30.999817	0.96	-13.123454	0.91873	-13.009308	0.909047	-16.369782	0.885
e	0.0009155	1	0.003884	0.663	0.009358	0.14675	0.009753	0.134504	0.002083	0.803
Z_1	12.31	0.985								
Z_2	12.31	0.985	6.391854	0.965	1.85907	< 2e-16	1.91767	< 2e-16	2.727711	< 2e-16
Z_3	12.17	0.987	6.297434	0.965	1.707293	< 2e-16	1.770031	< 2e-16	2.639204	2.55E-15
Z_4	12.3	0.985	6.287074	0.966	1.806892	< 2e-16	1.849606	< 2e-16	2.645964	< 2e-16
Z_5	-12.34	0.985	-6.264995	0.966					-2.624085	< 2e-16
Z_6	-11.79	0.986	-5.922376	0.968			-1.242759	0.000793		
Z_7	60.6	0.976	30.88341	0.961	10.813382	0.933	11.840088	0.917191	14.042288	0.901
Z_8	12.35	0.986	6.136427	0.966	1.534087	< 2e-16	1.554333	< 2e-16	2.570655	< 2e-16
Z_9	-0.01292	1	0.209967	0.206	0.348815	0.00376	0.357521	0.003381	0.168865	0.276

Variable no considerada en la corrida

Fuente: Elaboración propia

3.3.2 Validación de la aplicación del modelo

Tras realizar todas las ejecuciones del modelo Logit y Probit, por cada año y considerando la combinación de variables, es necesario elegir el modelo y la ecuación del Modelo Lineal Generalizado que ajusta con mayor precisión. Para tal fin, se consideran los criterios especificados en el *Diseño metodológico*.

La heterogeneidad de la información segmenta los datos en dos grupos: los años 2004 y 2006 en un grupo y la información para 2008, 2010 y 2012, en otro. En los apartados siguientes se analizan los resultados para determinar el modelo que mejor ajusta a los datos de acuerdo al grupo al que pertenecen.

Información para 2004 y 2006

Para 2004 y 2006 se ejecutaron tres corridas para cada modelo. La corrida inicial (cuadro 3.5) presenta la divergencia del modelo y por tanto se ejecutan corridas adicionales eliminando las variables donde la carencia de información repercute en el modelo. En el cuadro 3.16 se resaltan los resultados para la corrida 2, donde se selecciona el modelo logit como recomendado pues el porcentaje de explicación de la variabilidad es mayor y el AIC es el mejor entre ambos modelos.

Cuadro 3.16
Indicadores de validez de modelos aplicados – 2004
Indicadores por modelo y combinaciones de variables para determinar el modelo final

Variable	Corrida 1		Corrida 2	
	Logit	Probit	Logit	Probit
	p-value	p-value	p-value	p-value
β_1				
e	0.6881	0.6871	0.6881	0.6881
Z_2	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2.2e-16
Z_3	5.35E-11	5.351E-11	7.392E-12	7.392E-12
Z_4	2.02E-15	1.977E-15	2.02E-15	2.02E-15
Z_5	2.71E-07	2.816E-07		
Z_6				
Z_7	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2.2e-16
Z_8	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2.2e-16
Z_9	0.6002	0.6005	0.9538	0.9538
Porcentaje de explicación de la variabilidad	100.00%	100.00%	97.87%	97.39%
AIC	18	18	68.366	80.072
	Variable no considerada en la corrida			

Fuente: Elaboración propia

Para la estimación del 2006, se utilizan criterios idénticos a la estimación de 2004. Sin embargo, los p-values asociados a las variables e , Z_4 (rezago educativo), Z_8 (índice de marginación) y Z_9 (zona de residencia) son menores en el modelo probit, aunque el modelo logit es mejor de acuerdo al criterio del AIC y considerando el porcentaje de explicación de la variabilidad. Por lo tanto, el modelo recomendado para 2006 es el modelo Logit.

Cuadro 3.17
Indicadores de validez de modelos aplicados - 2006
Indicadores por modelo y combinaciones de variables para determinar el
modelo final

Variable	Corrida 1		Corrida 2	
	Logit	Probit	Logit	Probit
	p-value	p-value	p-value	p-value
β_1				
e	0.2945	0.2929	0.2945	0.2929
Z_2	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2.2e-16
Z_3	1.36E-14	1.36E-14	< 2.2e-16	< 2.2e-16
Z_4	3.649E-11	3.314E-11	3.649E-11	3.314E-11
Z_5	< 2.2e-16	< 2.2e-16		
Z_6				
Z_7	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2.2e-16
Z_8	< 2.2e-16	< 2.2e-16	5.566E-16	3.568E-13
Z_9	0.7024	0.7023	3.82E-01	3.16E-01
Porcentaje de explicación de la variabilidad	100.00%	100.00%	92.88%	92.42%
AIC	18	18	201.09	213.06

■ Variable no considerada en la corrida

Fuente: Elaboración propia

Información para 2008, 2010 y 2012

En las tablas 3.18, 3.19, 3.20, 3.21, 3.22 y 3.23 se remarcan los mejores indicadores por modelo y combinación de variables. Para todos los modelos previos a 2008 el modelo recomendado es el modelo Logit, aunque la diferencia entre éste y el modelo Probit es relativamente poca respecto a la explicación de la variabilidad. La corrida 1 y 2 de cada tabla mostrada no se consideran en la evaluación pues el modelo diverge y por tanto no es sujeto de validación.

Para el año 2008 el modelo se ajusta con mayor significancia para la corrida 5, al retirar las variables Z_1 (acceso a la alimentación) y Z_6 (acceso a los servicios médicos) donde el porcentaje de explicación de la variabilidad es mayor y el AIC es el menor (485.82).

Cuadro 3.18
Indicadores de validez de modelo Logit - 2008
Indicadores por modelo y combinaciones de variables para determinar el modelo final

Variable	Corrida 1	Corrida 2	Corrida 3	Corrida 4	Corrida 5
	p-value	p-value	p-value	p-value	p-value
β_1					
e	0.10	0.10	0.10	0.10	0.10
Z_1	< 2.2e-16				
Z_2	0.50	0.50	0.73	0.61	0.41
Z_3	< 2.2e-16	< 2.2e-16	0.00	< 2.2e-16	< 2.2e-16
Z_4	< 2.2e-16	< 2.2e-16	0.00	< 2.2e-16	< 2.2e-16
Z_5	< 2.2e-16	< 2.2e-16			< 2.2e-16
Z_6	0.00	0.00		0.00	
Z_7	< 2.2e-16	< 2.2e-16	0.00	< 2.2e-16	< 2.2e-16
Z_8	< 2.2e-16	< 2.2e-16	0.00	< 2.2e-16	< 2.2e-16
Z_9	0.00	0.00	0.00	0.00	0.00
Porcentaje de explicación de la variabilidad	100.00%	88.58%	79.37%	80.58%	86.66%
AIC	22.00	420.25	739.40	698.79	485.82

Variable no considerada en la corrida

Fuente: Elaboración propia

Análogamente, al analizar la información del modelo Probit para 2008, éste se aproxima mejor al retirar las variables Z_1 (acceso a alimentación) y Z_6 (acceso a servicios médicos) donde el porcentaje de explicación de la variabilidad es mayor y el AIC es el menor.

Cuadro 3.19
Indicadores de validez de modelo Probit - 2008
Indicadores por modelo y combinaciones de variables para determinar el modelo final

Variable	Corrida 1	Corrida 2	Corrida 3	Corrida 4	Corrida 5
	p-value	p-value	p-value	p-value	p-value
β_1					
e	0.10	0.10	0.10	0.10	0.10
Z_1	<2.2e-16				0.00
Z_2	0.50	0.50	0.68	0.57	0.41
Z_3	<2.2e-16	<2.2e-16	0.00	<2.2e-16	<2.2e-16
Z_4	<2.2e-16	<2.2e-16	0.00	<2.2e-16	<2.2e-16
Z_5	<2.2e-16	<2.2e-16			<2.2e-16
Z_6	0.00	0.00		0.00	

Indicadores por modelo y combinaciones de variables para determinar el modelo final

Variable	Corrida 1	Corrida 2	Corrida 3	Corrida 4	Corrida 5
	p-value	p-value	p-value	p-value	p-value
Z_7	<2.2e-16	<2.2e-16	0.00	<2.2e-16	<2.2e-16
Z_8	<2.2e-16	<2.2e-16	0.00	<2.2e-16	<2.2e-16
Z_9	0.00	0.00	0.00	0.00	<2.2e-16
Porcentaje de explicación de la variabilidad	100.00%	88.59%	79.50%	80.73%	86.72%
AIC	22.00	420.20	734.85	693.69	483.75

Variable no considerada en la corrida

Fuente: Elaboración propia

Para el año 2010, el AIC utilizando el modelo probit es menor que el AIC para el modelo Logit, y solo en una variable se diferencia el modelo, Z_9 variable en la que el modelo Logit aproxima adecuadamente. Sin embargo, en el aspecto global, el modelo Probit debe ser utilizado para estimar.

Cuadro3.20

Indicadores de validez del modelo logit-2010

Indicadores por modelo y combinaciones de variables para determinar el modelo final

Variable	Corrida 1	Corrida 2	Corrida 3	Corrida 4	Corrida 5
	p-value	p-value	p-value	p-value	p-value
β_1					
e	0.45172	0.4517	0.45172	0.451722	0.4517
Z_1	2.20E-16				
Z_2	2.20E-16	< 2.2e-16	< 2e-16	< 2.2e-16	<2e-16
Z_3	2.20E-16	< 2.2e-16	< 2e-16	< 2.2e-16	<2e-16
Z_4	2.20E-16	< 2.2e-16	< 2e-16	< 2.2e-16	<2e-16
Z_5	2.20E-16	< 2.2e-16			<2e-16
Z_6	1.74E-10	2.17E-09		2.17E-09	
Z_7	2.20E-16	< 2.2e-16	< 2e-16	< 2.2e-16	<2e-16
Z_8	2.20E-16	< 2.2e-16	< 2e-16	< 2.2e-16	<2e-16
Z_9	0.04344	0.2825	0.01242	0.004617	0.4373
Porcentaje de explicación de la variabilidad	100.00%	87.97%	78.95%	80.30%	86.31%
AIC	22.00	594.26	1,021.10	958.38	671.75

Variable no considerada en la corrida

Fuente: Elaboración propia

Para 2010, el modelo Probit resulta ser el más adecuado considerando los antecedentes y por la proporción de significancia del modelo y por la aportación de la variabilidad.

Cuadro3.21
Indicadores de validez del modelo Probit-2010
Indicadores por modelo y combinaciones de variables para determinar el modelo final

Variable	Corrida 1	Corrida 2	Corrida 3	Corrida 4	Corrida 5
	p-value	p-value	p-value	p-value	p-value
β_1					
e	0.45062	0.4506	0.4506	0.450622	0.4506
Z_1	2.20E-16				
Z_2	2.20E-16	2.20E-16	<2e-16	< 2.2e-16	<2e-16
Z_3	2.20E-16	2.20E-16	<2e-16	< 2.2e-16	<2e-16
Z_4	2.20E-16	2.20E-16	<2e-16	< 2.2e-16	<2e-16
Z_5	2.20E-16	2.20E-16			<2e-16
Z_6	5.42E-10	5.42E-10		1.54E-09	
Z_7	2.20E-16	2.20E-16	<2e-16	< 2.2e-16	<2e-16
Z_8	2.20E-16	2.20E-16	<2e-16	< 2.2e-16	<2e-16
Z_9	0.04316	0.3003	0.0131	0.004938	0.4546
Porcentaje de explicación de la variabilidad	100.00%	87.97%	78.99%	80.42%	86.36%
AIC	22.00	594.15	1,019.00	953.04	669.03

Variable no considerada en la corrida

Fuente: Elaboración propia

Los resultados del modelo Logit son similares a los del modelo Probit para 2012, por lo que se le trata con una serie de datos similares. Por lo tanto, el modelo recomendado es el Probit.

Cuadro3.22
Indicadores de validez del modelo Logit-2012
Indicadores por modelo y combinaciones de variables para determinar el modelo final

Variable	Corrida 1	Corrida 2	Corrida 3	Corrida 4	Corrida 5
	p-value	p-value	p-value	p-value	p-value
β_1					
e	0.010762	0.010762	0.010762	0.010762	0.01076
Z_1	< 2.2e-16				
Z_2	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2e-16
Z_3	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2e-16
Z_4	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2e-16
Z_5	< 2.2e-16	< 2.2e-16			< 2e-16
Z_6	0.007835	0.007835		0.007835	
Z_7	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2e-16

Indicadores por modelo y combinaciones de variables para determinar el modelo final

Variable	Corrida 1	Corrida 2	Corrida 3	Corrida 4	Corrida 5
	p-value	p-value	p-value	p-value	p-value
Z_8	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2e-16
Z_9	0.002277	0.00802	0.000152	0.000178	0.01158
Porcentaje de explicación de la variabilidad	100.00%	89.15%	80.73%	81.14%	88.11%
AIC	22.00	351.43	604.90	594.27	381.30

Variable no considerada en la corrida

Fuente: Elaboración propia

Para 2012, la situación es similar sin embargo es necesario especificar las variables que se encuentran expuestas para así determinar un modelo. Tal como se presenta en el cuadro 3.23, el modelo que mejor ajusta es el asociado a la corrida 5.

Cuadro3.23

Indicadores de validez del modelo Probit-2012

Indicadores por modelo y combinaciones de variables para determinar el modelo final

Variable	Corrida 1	Corrida 2	Corrida 3	Corrida 4	Corrida 5
	p-value	p-value	p-value	p-value	p-value
β_1					
e	0.010764	0.010764	0.010764	0.010764	0.01076
Z_1	< 2.2e-16				
Z_2	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2e-16
Z_3	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2e-16
Z_4	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2e-16
Z_5	< 2.2e-16	< 2.2e-16			< 2e-16
Z_6	0.011363	0.011363		0.011363	
Z_7	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2e-16
Z_8	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2e-16
Z_9	0.002402	0.007829	0.0001061	0.0001027	0.01023
Porcentaje de explicación de la variabilidad	100.00%	89.16%	80.83%	81.27%	88.12%
AIC	22.00	351.40	601.88	590.30	380.97

Variable no considerada en la corrida

Fuente: Elaboración propia

Para los modelos creados para los años 2008, 2010 y 2012, el mejor modelo resulta ser el Probit con la combinación de variables que elimina exclusivamente la variable Z_6 asociada al acceso a los servicios médicos.

3.3.3 Ajuste del modelo

El ajuste del modelo se realiza conforme al progreso del proyecto, por lo que en conclusión se tienen los Modelos Lineales Generalizados con el mejor ajuste por año:

Cuadro 3.24
Coefficientes finales de la modelación
Coefficientes obtenidos por cada modelo Probit

Variable	Predictor lineal	Año				
		2004	2006	2008	2010	2012
β_1	β_1	-28.84	-28.84	-14.62	-15.284084	-16.369782
e	β_2	0.004982	0.004982	0.0009954	-0.008315	0.002083
Z_1	β_3					
Z_2	β_4	7.984	7.984	1.743	2.375814	2.727711
Z_3	β_5	7.652	7.652	2.856	2.787471	2.639204
Z_4	β_6	3.39	3.39	2.212	2.514885	2.645964
Z_5	β_7			-2.304	-2.425898	-2.624085
Z_6	β_8					
Z_7	β_9	26.75	26.75	13.14	13.491614	14.042288
Z_8	β_{10}	3.082	3.082	2.099	2.286703	2.570655
Z_9	β_{11}	0.5211	0.5211	-0.04804	-0.247397	0.168865

Fuente: Elaboración propia

Entonces, los coeficientes presentados por año aplicados a la ecuación 18 presentada en el capítulo dos, proporciona el valor de la probabilidad π_i . De los coeficientes presentados en el cuadro 3.24 se reafirma que algunas variables no resultan relevantes para el estudio de la pobreza extrema y desvirtuando el modelo de ajuste.

3.4 Resultados

Parte del objetivo de este proyecto es presentar el modelo y los resultados que arroja, consecuencia de la estimación de parámetros realizada mediante el Modelo Lineal Generalizado. En éste apartado se presentan los resultados obtenidos con los coeficientes presentados en el cuadro 3.24 y aplicando la función liga correspondiente. Los resultados mostrados en este apartado son comparables respecto a los cuadros presentados en el apartado 1.3 *Análisis cuantitativo de la pobreza*.

Finalmente para obtener las proporciones finales se aplica el Teorema de Probabilidad Total estratificado a nivel nacional o por estado.

Nivel nacional

Para calcular la proporción de la población en situación de pobreza, se aplica el Teorema de Probabilidad total a los predictores lineales obtenidos y a las probabilidades de cada combinación de carencias. De dicho cálculo resulta la matriz de proporciones por año a nivel nacional. La información para 2008 y 2010 es comparable con el cuadro 1.2 del capítulo 1.

La información respecto al cuadro 1.2 y el cuadro 3.25 es distinta pues la manipulación de los datos y el modelo utilizado proporciona diferencias, que difícilmente pueden subsanarse.

Cuadro 3.25
Comparación de la proporción de la población en situación de pobreza extrema de acuerdo al modelo lineal generalizado respecto a estimación CONEVAL

Año	Proporción de la población en pobreza extrema		
	Modelo lineal generalizado	Estimaciones CONEVAL	Diferencia
2004	4.58%		
2006	6.47%		
2008	3.88%	10.60%	6.72%
2010	3.02%	10.40%	7.38%
2012	5.23%		

Fuente: Elaboración propia

En el cuadro 3.25 se presentan las diferencias respecto al cuadro 1.2 para 2008 y 2010 siendo esta información la disponible, pues la primera medición mutidimensional de la pobreza se realizó en 2008 y los resultados para 2012 se concluyen y publican posterior a la elaboración de este trabajo.

El método de medida del CONEVAL comprende el conteo general de carencias y no vincula la influencia de las variables. Por lo tanto, existe una diferencia de 6.72% y 7.38% para 2008 y 2010 respectivamente.

Conclusiones

El mundo reconoce a México como un país de gente con creatividad, que siempre encuentra la forma de mejorar su situación. Sin embargo es también un país de contrastes en todos los sentidos de la palabra y más aún hablando de distribución de la riqueza, pues entre los habitantes se encuentra uno de los hombres más adinerados del mundo, también existe una tasa de pobreza muy alta.

Para mejorar y retirar la calidad de país subdesarrollado, se deben enfocar los esfuerzos en mitigar o erradicar los factores que perpetúan a la pobreza y de los cuales, el mexicano promedio vive día a día.

En el capítulo uno se esboza la pobreza cómo un problema con muchos matices, que se ha estudiado desde distintos enfoques hasta concluir, para el caso mexicano, en la medición de la pobreza multidimensional con una lista de indicadores mínimos a considerar como obligatorios tras la promulgación de la Ley General de Desarrollo Social.

La modificación responde a los esfuerzos internacionales por mejorar la calidad de vida de los habitantes y ayudar a los países subdesarrollados a prosperar, pues considera que la población debe cubrir sus necesidades básicas como son alimentación, vivienda y educación.

Para enfocar cualquier iniciativa es necesario medir cualquier acción ejecutada para aprender y mejorar la situación. Concretamente, la medición de cualquier fenómeno se dificulta si no se cuenta con información suficiente, oportuna y confiable; en el caso de la pobreza el tema se complica pues, tal como describen algunos autores, la pobreza también es relativa al contexto social, cultural o político en el que se encuentre el investigador y por lo tanto, debe crearse una escala cuyo uso sea factible a lo largo de un determinado período.

Los problemas citados se permean en la creación de este trabajo, pues el cambio en una política pública convertida en ley, modifica el parámetro de medida, tal como ocurrido al modificar la Constitución incluyendo la educación media superior como parte de la educación básica y obligatoria.

La promulgación de la Ley de Desarrollo Social indujo cambios de raíz, para así garantizar el cumplimiento de la misma. Los cambios se dieron a todos los niveles: creación del CONEVAL, modificación de la estructura para la atención de la pobreza, la inclusión de la metodología multidimensional en la medición de la pobreza, cambios en la recolección de información para medir la pobreza, entre otras.

Consecuencia de los cambios, se requirió cambiar paulatinamente la estructura de las bases de datos, las modificaciones en catálogos, los cambios en nombre de archivos, la inclusión y eliminación de campos, entre otros; sin embargo, estas modificaciones complicaron la manipulación de las bases para la elaboración de este trabajo, sobre todo al homologarlas. El objeto de la homologación es tener unidades de medida comparables entre los distintos años de levantamiento.

La información utilizada considera una serie de modificaciones y tratamientos, en todos los años estudiados, que se reflejan en el uso de exclusivamente algunos campos de la Encuesta Nacional de Ingresos y Gastos en Hogares, de tal forma que los campos efectivamente aportan al modelo a estudiar de acuerdo a las variables definidas en la Ley.

Dadas las características de la información disponible, ésta se divide en dos grupos para su manipulación: la correspondiente a 2004 y 2006 y, en un segundo grupo la información de 2008, 2010 y 2012. La separación atiende a una modificación en la estructura de la base de datos donde el grupo 1 (2004 y 2006) tiene estructura idéntica entre sí y los cambios son ínfimos, consecuencia de la inexistencia de la Ley. Para el segundo grupo 2 (bases posteriores a 2008), la manipulación se complica dado que la modificación en la Ley considera un cambio para recabar más datos y por lo tanto se incluyen campos, se modifican relaciones entre tablas, etc., que complica la homologación de las bases.

El objetivo es medir la pobreza extrema por lo que la categorización de la población resulta en: pobre extremo o no pobre extremo, donde la ausencia de pobreza extrema incluye todas las variantes contempladas en los Lineamientos publicados por el CONEVAL.

Por lo tanto, en este proyecto se utiliza un modelo de regresión logística dicotómico, de acuerdo a la recomendación de CEPAL y UNICEF para mediciones de pobreza infantil. También es factible el un modelo politómico, pero éste induce una gran cantidad de información y estratificación de los resultados a realizarse al considerar una variable respuesta con más de una opción.

Entonces, dado el tratamiento de la información en dos grupos, se tiene de igual manera que el modelo Logit es el que mejor ajusta para la información previa a 2008. Sin embargo, para la información de 2008 y posterior, es el modelo Probit el que mejor estima.

El modelo es ajustado después de analizar la información y concluir que no todas las variables son requeridas para garantizar la convergencia. Se retiraron las variables que miden el acceso a la alimentación y el acceso a los servicios médicos. La variable asociada al alimentación se retiró en el modelo de todos los años, por la carencia de la información o

bien, por la poca relevancia que esta presenta, consecuencia de la ambigüedad de las preguntas en las que más bien se mide percepción.

Finalmente, para el modelo de 2004 las variables que explican y aportan al modelo son *servicios en la vivienda y rezago educativo*. Para 2006 el resultado es similar, añadiéndose la variable *índice de marginación*, asociado a una estimación de CONAPO. Para 2008, 2010 y 2012, las variables significativas son *calidad y servicios en la vivienda, rezago educativo, acceso a la seguridad social e índice de marginación*.

Con lo anterior queda demostrado que no todas las variables presentadas por CONEVAL son relevantes para la medición de la pobreza extrema, asumiendo los actuales estándares en el levantamiento de la información.

Dado que en los cuestionarios, la recolección de algunos indicadores se realiza como una medida de percepción, es preciso revisar las preguntas en los cuestionarios a modo de medida exacta para la información que se pretende recabar. De mucha utilidad resultaría considerar una validación de los datos para que la información contenida en las bases sea congruente con la realidad y tenga consistencia.

Los resultados obtenidos al aplicar el modelo resultaron ser diferentes a lo publicado por el CONEVAL, con diferencias de 6.72 y 7.38 puntos porcentuales respecto a la medición realizada para 2008 y 2010. Estas diferencias se deben al uso de un modelo distinto, la estratificación realizada para la homologación de la información y la modificación de acuerdo a las políticas vigentes a la fecha de elaboración de este trabajo.

El objetivo del estudio se cumplió al encontrar un modelo que estima el índice de la pobreza extrema, pues bajo las premisas establecidas por la Ley General de Desarrollo Social y el CONEVAL, y resulta factible utilizar un Modelo Lineal Generalizado. Sin embargo, la falta de madurez en la información arroja resultados poco consistentes comprobando que no todas las variables indicadas por los Lineamientos aportan información relevante a la medición de la pobreza extrema.

Finalmente, la complicación y reto más grande es definir la pobreza para de acuerdo a ella medirla. Pese a los esfuerzos de muchos investigadores e instituciones, el factor *contexto social* determina en gran medida la definición de *pobre* y este cambio considera variaciones en la medición.

Referencias Bibliográficas

De Groot, Morris H., *Probabilidad y estadística*, Addison – Wesley Iberoamericana, Segunda edición, 1988, p323-329.

Dobson, Annete J., *An introduction to generalized linear models*, Chapman and Hall/CRC, Segunda Edición, 2001, Capítulos 7 y 8.

Fox, John, *Applied Regression Analysis and Generalized Linear Models*, Sage Publications, Segunda Edición 2008, páginas 379-424.

Rincón, Luis, *Introducción a la probabilidad*, Facultad de Ciencias , agosto 2013, página 65

McCullagh, P. and J. A. Nelder, *Generalized Linear Models*, Chapman and Hall/CRC, Segunda Edición, 1989, páginas: 2-44,98-236

Smith, Adam, *Investigación de la naturaleza y causas de la riqueza de las naciones (1794)*, Edición digital de la Universidad de Sevilla.

Venables, W.N. and Ripley, B.D., *Modern Applied Statistics with S*, Springer, Cuarta edición, 2002, p.188

Amarya K. Sen , *Sobre conceptos y medidas de pobreza* , Comercio Exterior Vol. 42, Núm. 4, abril 1992.

Bartle, Phil trad. Sada Ma. Loudes, *Los cinco principales factores de la pobreza. Factores, no síntomas*, <http://cec.vcn.bc.ca/mpfc/modules/emp-povs.htm> (en español), Community Empowerment Collective. Consultado el 13 de julio, 2013.

Damian, Araceli y Boltvinik, Julio, *Evolución y características de la pobreza en México*, Comercio Exterior, Vol. 53, Núm. 6, junio 2003

Guisan, Antoine, Edwards, Thomas C., Hastie, Trevor, *Generalized linear al generalized additive models in studies of species distributions: setting the scene*, Ecological Modelling, Vol. 157, 2002, p. 89-100.

Nelder J.A. and Wedderburn, W.M., *Generalized Linear Models*, Journal of the Royal Statistical Society, Series A(General), Wiley for the Royal Society, Vol 135 No. 3, 1972, p. 370-384.

P.Burnham, Kenneth and Anderson,David R ,*Multimodel inference*, Sociological Methods & Research, Vol. 33, Núm. 2, noviembre 2004, p.261-304.

Schervish MJ, *P Values: What They Are and What They Are Not*. The American Statistician, Num 50 , vol.3,1996,p. 203-206.

Comisión económica para América Latina y el Caribe, Fondo de las Naciones Unidas para la Infancia, *Pobreza infantil en América Latina y el Caribe*, 2010, p.165-168.

Consejo Nacional de Evaluación de la Política de Desarrollo social, *Lineamientos y criterios generales para la definición, identificación y medición de la pobreza*, Ciudad de México, Diario Oficial de la Federación, 2010.

Consejo Nacional de Evaluación de la Política de Desarrollo social, *Informe de la Evaluación de la Política de Desarrollo Social en México 2010*, México, D.F. , CONEVAL, 2012.

Consejo Nacional de Evaluación de la Política de Desarrollo social, *Glosario*, <http://web.CONEVAL.gob.mx/Medicion/Paginas/Glosario.aspx> (en español), Consejo Nacional de Evaluación de la Política de Desarrollo Social. Consultado el 17 de julio, 2013.

Ley General de Desarrollo Social, Ciudad de México, Diario Oficial de la Federación, 16 de junio 2012.

Constitución Política de los Estados Unidos Mexicanos, Ciudad de México, Diario Oficial de la Federación, 5 de febrero de 1917, Última reforma publicada 26 de febrero, 2013.

Plan nacional de desarrollo 2007 - 2012, Ciudad de México, Gobierno de los Estados Unidos Mexicanos, mayo 2007, p.145.

Glosario

Término	Significado
CEPAL	Comisión Económica para América Latina y el Caribe
CNDH	Comisión Nacional de los Derechos Humanos
CONAPO	Consejo Nacional de Población
CONEVAL	Consejo Nacional de Evaluación de la Política de Desarrollo Social
DICONSA	Empresa de participación estatal mayoritaria que pertenece al Sector Desarrollo Social. Tiene el propósito de contribuir a la superación de la pobreza alimentaria, mediante el abasto de productos básicos y complementarios a localidades rurales de alta y muy alta marginación, con base en la organización y la participación comunitaria.
ENIGH	Encuesta Nacional de Ingresos y Gastos en Hogares
IGAE	Indicador Global de la Actividad Económica
IMSS	Instituto Mexicano del Seguro Social
INEGI	Instituto Nacional de Estadística y Geografía
ISSFAM	Instituto de Seguridad Social para las Fuerzas Armadas de México
ISSSTE	Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado.
LICONSA	Empresa de participación estatal mayoritaria, industrializa leche de elevada calidad y la distribuye a precio subsidiado en apoyo a la nutrición de millones de mexicanos, especialmente de niños de hasta 12 años, de familias en condiciones de pobreza, contribuyendo así a su adecuada incorporación al desarrollo del país.
PIB	Producto Interno Bruto
SEDESOL	Secretaría de Desarrollo Social
SEPOMEX	Servicio Postal Mexicano