



**UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MÉXICO**

---

---

**FACULTAD DE CIENCIAS**

**MODELO ESTADÍSTICO PARA LA  
DETERMINACIÓN DE FACTORES  
FAVORABLES PARA EL DISEÑO DE POLÍTICAS  
PÚBLICAS EN CIENCIA Y TECNOLOGÍA**

**T E S I S**

**QUE PARA OBTENER EL TÍTULO DE:**

**A C T U A R I O  
P R E S E N T A:**

**WILFRIDO URUETA RICO**



**DIRECTOR DE TESIS:  
DR. ENRIQUE RUIZ-VELASCO SÁNCHEZ  
CIUDAD DE MÉXICO, 2018**



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 del alumno**

Apellido paterno: Urueta  
Apellido materno: Rico  
Nombre(s): Wilfrido  
Teléfono: 15 39 15 54  
Universidad: Universidad Nacional Autónoma de México  
Escuela o facultad: Facultad de Ciencias  
Carrera: Actuaría  
Número de cuenta 086519579

**2. Datos del tutor**

Grado: Dr  
Nombre(s): Enrique  
Apellido paterno: Ruiz-Velasco  
Apellido materno: Sánchez

**3. Datos del sinodal 1**

Grado: Act.  
Nombre(s): Francisco  
Apellido paterno: Sánchez  
Apellido materno: Villarreal

**4. Datos del sinodal 2**

Grado: M en I  
Nombre(s): Karla  
Apellido paterno: Ramírez  
Apellido materno: Pulido

**5. Datos del sinodal 3**

Grado: M. en E.  
Nombre(s): Octavio Daniel  
Apellido paterno: Ríos  
Apellido materno: Lázaro

**6. Datos del sinodal 4**

Grado: Act.  
Nombre(s): Ma. Susana  
Apellido paterno: Barrera  
Apellido materno: Ocampo

**7. Datos del trabajo escrito.**

Título: Modelo estadístico para la determinación de políticas públicas en ciencia y tecnología

Subtítulo:  
Número de páginas: 72  
Año: 2018

## *Agradecimientos*

*A mis maestros Karla Ramírez, Susana Barrera, Octavio Ríos, Francisco Sánchez Villarreal y Enrique Ruiz-Velasco por su orientación profesional y sobre todo por su magnífico apoyo para realizar correctamente este trabajo. Todos y cada uno de ustedes me hacen sentir honrado.*

*A mis amigos Ricardo Blanco y Octavio Ríos, con quienes he compartido muchos años de trabajo y amistad y han sido promotores incansables de mi titulación. Les agradezco mucho todo ese apoyo y la confianza depositada en mí.*

*A Loredmy por apoyarme decididamente en este proceso, así como por muchas alegrías compartidas durante años, pero sobre todo por haberme dado mis mayores tesoros: mis hijos amados.*

## *Dedicatorias*

*A mi madre Lucía (†) y a mi padre Wilfrido (†), quienes hicieron de mi y de mis hermanos personas con los más altos valores éticos. ¡Lamento mucho no haberles dado esta satisfacción en vida!*

*A mis hermanos fallecidos Jacobo, Eréndira, Alfonso y Jesús, a quienes extraño mucho.*

*A mi hermana Juanita, que además es como mi segunda madre y a mi cuñado Sigfrido, quien es como mi hermano.*

*A mi hermano Cuauhtémoc, en quien mi padre depositó su confianza para impulsarme en la vida en su representación y a Araceli, quien le dio una renovada vida a mi hermano.*

*A mi hermano Gilberto, quien siempre ha sido mi gran amigo.*

*A mis hermanos Osbelia, Jair y Alma Rosa con mucho cariño.*

*A mis primos Jaime, Eduardo, Gabriela, Miguel y David, que son como mis hermanos.*

*A mis sobrinos Pavel, Sigfrido, Adonahi y Alfonso, que son como mis hermanitos menores y con los que he pasado grandes momentos.*

*A mis otros sobrinos Tathali, Innan, Arturo, Cuiclahuac, Christian, Abril, Xóchitl y Omar, que son mi familia querida.*

*A mis amados hijos Loredmy y Wilfrido, quienes han venido a iluminar mi vida.*

*“Vivimos en una sociedad profundamente dependiente de la ciencia y la tecnología y en la que nadie sabe nada de estos temas. Ello constituye una fórmula segura para el desastre.”*

Carl Sagan 1934-1996. Astrónomo estadounidense.

## ÍNDICE GENERAL

<b>OBJETIVOS</b>	8
<b>HIPÓTESIS</b>	8
<b>I. INTRODUCCIÓN</b>	9
1.1 Bienestar nacional derivado de la investigación y desarrollo	9
1.2 Ciencia y tecnología, percepción pública y cultura científica	11
1.3 Antecedentes	14
1.3.1 Primer estudio de análisis de trayectorias	15
1.3.2 Estudio de análisis de trayectorias en México	17
<b>II FUENTE DE INFORMACIÓN</b>	20
2.1 Encuesta sobre la Percepción Pública de la Ciencia y la Tecnología en México, ENPECYT 2013	20
2.1.1 Objetivo general de la ENPECYT	20
2.1.2 Cobertura de la encuesta	21
2.1.3 Unidad de observación	21
2.1.4 Diseño del cuestionario	22
2.1.5 Esquema de muestreo	22
2.1.6 Marco muestral	22
2.1.7 Estratificación	23
2.1.8 Tamaño de la muestra	23
2.1.9 Afijación de la muestra	24
2.1.10 Selección de la muestra	25
2.1.11 Cálculo de los factores de expansión	25
2.1.12 Ajustes a los factores de expansión	26
2.1.13 Estimadores	27
2.1.14 Estimaciones de errores de muestreo	27
2.2 Observaciones al diseño muestral	29
<b>III DESARROLLO DEL TEMA</b>	31
3.1 Generalidades	31
3.2 Modelo de cultura científica del Manual de Antigua	31
3.3 Crítica al modelo del Manual de Antigua	33
3.4 Modelo alternativo de cultura científica	33
3.5 Modelo conceptual y estadístico	34
3.5.1 Variables del Modelo conceptual y estadístico	36
<b>IV REGRESIÓN LOGÍSTICA MULTIVARIANTE (MODELO LOGIT)</b>	40

4.1	Introducción	40
4.2	Modelo de Regresión Logística en relación con el Modelo de Regresión Lineal Múltiple	41
4.3	Planteamiento del Modelo de Regresión Logística	42
4.3.1	Regresión logística	44
4.3.2	Estimación del modelo logístico	45
4.3.3	Ajuste del Modelo Logit	47
4.3.4	Grado de contribución de las variables de forma individual	49
4.3.5	Interpretación de los coeficientes del modelo	49
<b>V RESULTADOS ESTADÍSTICOS</b>		<b>51</b>
5.1	Resultados de los ejercicios estadísticos	51
5.2	Modelo resultante	53
5.3	Análisis de los resultados	55
<b>VI CONCLUSIONES Y RECOMENDACIONES</b>		<b>58</b>
6.1	Conclusiones	58
6.2	Cumplimiento de los objetivos	58
6.3	Cumplimiento de las hipótesis	59
6.4	Recomendaciones de políticas públicas	62
<b>ANEXO 1. EJERCICIO ESTADÍSTICO CON 25 VARIABLES</b>		<b>64</b>
<b>ANEXO 2. EJERCICIO ESTADÍSTICO DEFINITIVO</b>		<b>68</b>
<b>BIBLIOGRAFÍA</b>		<b>71</b>

## ÍNDICE DE TABLAS

Tabla 1.	GIDE/PIB por país, 2016 [IGECTI 2016] [IMD]	10
Tabla 2.	Clasificación de las variables y sus escalas de medición	17
Tabla 3.	Distribución de la muestra de viviendas por área urbana	24
Tabla 4.	Distribución de la tasa de no respuesta observada en los ejercicios de la ENPECYT	29
Tabla 5.	Construcción de dos variables binarias a partir de una variable categórica de 3 valores excluyentes	41
Tabla 6.	Esquema de la Tabla de Clasificación	48
Tabla 7.	Tabla de clasificación considerando 25 variables dependientes	51
Tabla 8.	Resultados estadísticos del modelo Logit con 25 variables dependientes	52
Tabla 9	Tabla de clasificación del modelo resultante	53

Tabla 10. Resultados estadísticos del modelo resultante	54
Tabla 11. Probabilidades individuales y momios asociados a cada variable	55
Tabla 12. Percepción del grado de contenido científico de diferentes disciplinas en México, 2013	61

## ÍNDICE DE FIGURAS

Figura 1 Esquema de Análisis de Trayectorias (ejemplo ilustrativo)	16
Figura 2. Análisis de Trayectorias de México, 1999	19
Figura 3. Modelo de Escalera de Cultura Científica del Manual de Antigua, 2012	32
Figura 4. Cultura científica, Esquema alternativo	34
Figura 5. Diagrama del Modelo conceptual y estadístico	35
Figura 6. Esquema del Modelo de Regresión Logística Multivariante	40
Figura 7. Interpretación geométrica del método de Newton-Rapson	47

## **OBJETIVOS DE LA TESIS**

1. Encontrar el conjunto de variables o factores que incidan en el apoyo de los ciudadanos a las políticas públicas en ciencia y tecnología.
2. Identificar variables o factores específicos y su impacto individual en el apoyo a las políticas públicas en ciencia y tecnología.
3. Medir la probabilidad de apoyo a políticas públicas en ciencia y tecnología, dado un caso particular de ciudadano con perfil específico.
4. Emitir recomendaciones para fortalecer o definir políticas públicas en favor de la ciencia y la tecnología, con base en los resultados del ejercicio estadístico.

## **HIPÓTESIS**

**Hipótesis 1:** La formación educativa es un factor relevante en la percepción pública y en el apoyo a las políticas públicas en ciencia y tecnología.

**Hipótesis 2:** La consulta regular de información científica y tecnológica de las personas es un factor que impacta fuertemente en el apoyo a las políticas públicas en ciencia y tecnología.

# I INTRODUCCIÓN

En este capítulo se presentará una descripción general de la importancia que tiene la ciencia y tecnología en el desarrollo económico de los países, así como su relación con la percepción y cultura científica y tecnológica de las personas, y organismos públicos y privados, como factor favorable al impulso de políticas públicas en esta materia. Asimismo, se presentan los conceptos que se abordarán en esta tesis y sus definiciones y clasificaciones homologadas a nivel internacional. Finalmente, se reportarán los estudios de orden predictivo que precedieron al desarrollado en esta tesis, tanto a nivel mundial como en México.

## 1.1 Bienestar nacional derivado de la investigación y desarrollo.

Un país que no invierte en investigación científica y desarrollo tecnológico está destinado a ser dependiente de los que desarrollan su propio conocimiento y tecnología, debido a que no genera los elementos de competitividad nacional e internacional.

Para justificar lo mencionado en el párrafo anterior, cabe mencionar que el Foro Económico Mundial, a través del Instituto para el Desarrollo Empresarial (IMD por sus siglas en inglés), presenta de forma anual los indicadores de competitividad para países seleccionados [8]. El principal indicador del IMD es el índice de competitividad, el cual permite clasificar a cada país en su posición correspondiente.

El índice de competitividad se compone de cuatro grandes rubros, que son: a) desempeño económico del país, el cual está basado en la evaluación macroeconómica de la economía doméstica; b) eficiencia gubernamental, explicada por el grado en que las políticas del gobierno conducen a la competitividad; c) eficiencia empresarial, que es el grado en el que las empresas se desempeñan de manera innovadora, rentable y responsable; y d) infraestructura, la cual se caracteriza por el grado en que los recursos productivos, tecnológicos, científicos y humanos satisfacen las necesidades de las empresas. En cada uno de los rubros mencionados hay componentes de ciencia y tecnología que abonan a su cálculo.

De lo anterior, se puede inferir que un país que es dependiente económicamente de otros se caracteriza por su debilidad en el desempeño de uno o más de los factores que explican su competitividad. Cuando ello ocurre, los eventos internacionales de comercio, de finanzas o de política suelen afectarlos de tal manera que se ven obligados a devaluar su moneda, por ejemplo, para poder competir en el comercio internacional con los países cuyas empresas fabrican productos con mayor valor agregado con base en su desarrollo tecnológico.

Por otro lado, el principal indicador para medir de forma directa los esfuerzos agregados en el avance del conocimiento científico y el desarrollo de nuevos

productos y tecnologías es el gasto en investigación y desarrollo experimental en relación con el producto interno bruto (GIDE/PIB) [4].

**Tabla 1. GIDE/PIB por país, 2016 [4] [8]**

<b>País</b>	<b>Competitividad posición</b>	<b>GIDE/PIB %</b>	<b>PIB per cápita dólares de EUA</b>
Suiza	2	3.42	63,888.7
Estados Unidos	3	2.79	57,591.2
Singapur	4	2.18	52,960.7
Suecia	5	3.31	48,639.2
Canadá	10	1.67	44,819.5
Alemania	12	2.93	49,187.3
Finlandia	20	2.90	43,446.1
Israel	21	4.15	37,460.9
China	25	2.07	15,478.0
Japón	26	3.29	42,292.7
Corea	29	4.23	36,629.7
Francia	32	2.22	41,357.8
España	34	1.24	36,338.7
Chile	36	0.38	23,003.9
<b>México</b>	<b>45</b>	<b>0.50</b>	<b>18,535.2</b>
Argentina	55	0.63	20,382.4
Brasil	57.00	1.14	15,224.5

De acuerdo con la tabla anterior, se puede apreciar que los países que invierten un porcentaje mayor de su PIB en investigación y desarrollo (IDE) también son los reconocidos como más avanzados económicamente y mejor posicionados competitivamente, así como también con mejores logros sociales, en varios casos.

En los primeros cuatro países de la lista, es decir, Suiza, Estados Unidos, Singapur y Suecia, que son los más competitivos, se puede apreciar que su GIDE/PIB rebasa claramente el 2% y que su PIB per cápita es muy alto, por encima de 48 mil dólares. En el polo opuesto se encuentran los países latinoamericanos, cuyo GIDE/PIB es menor a 1%, salvo Brasil que invirtió 1.14 en ese rubro. Tales países están clasificados por debajo de la posición 36 en competitividad y sus niveles de PIB per cápita son claramente los más bajos observados en la tabla, salvo el caso de China.

Así, cuando un país democrático invierte decididamente en estos temas, su población, empresas e instituciones se benefician en conjunto, reportando mayor nivel y calidad de vida para la población, mediante trabajos mejor remunerados, más altos niveles educativos y mejores accesos a nuevas tecnologías, entre otros; por su parte, las empresas mejoran su competitividad y con ello se posicionan mejor en

los mercados, y estos resultados junto con otros factores impulsan el crecimiento económico y social del país en cuestión.

Para que un país alcance una situación como la planteada anteriormente, es menester que una proporción amplia de la población, incluyendo a empresarios y funcionarios públicos, incremente su cultura científica y tecnológica, de manera que vean a este tema con admiración y como un objetivo a alcanzar [3]. La falta de cultura científica de la población reduce la importancia de este tema y no hace posible la participación ciudadana en discusiones sobre políticas públicas de ciencia y tecnología [13]. Asimismo, la débil cultura científica y tecnológica es un obstáculo para las inversiones privadas en innovaciones tecnológicas con alto valor agregado y, por lo mismo, genera una posición tibia del gobierno para favorecer al sector ciencia y tecnología, relegándolo, como es el caso de México, a un segundo plano en la agenda gubernamental.

## **1.2 Ciencia y tecnología, percepción pública y cultura científica.**

Al igual que la cultura desarrollada en temas diversos como política, deportes, medio ambiente, educación o artes, entre otros, la *cultura científica y tecnológica* de las personas en lo individual, y de la sociedad en general, es resultado y a la vez causa de la combinación de diversos factores que, a través del tiempo, inciden en la adquisición y ampliación de conocimientos en el tema que, a su vez, implican percepciones y opiniones en torno a la importancia y el impacto de estas actividades, impactando en la postura favorable (o contraria) al avance en el conocimiento científico y el desarrollo tecnológico.

A continuación, se ofrecen las definiciones de los principales conceptos utilizados en la presente tesis.

Tanto la ciencia como la tecnología son conceptos abstractos que, a través del tiempo, han sido definidos de diferentes formas y bajo diversos criterios y clasificaciones, sin que se haya llegado a un consenso.

En el Diccionario de la Real Academia Española [21], se ofrecen las definiciones de cada uno de los términos:

### ***Ciencia:***

“1. f. Conjunto de conocimientos obtenidos mediante la observación y el razonamiento, sistemáticamente estructurados y de los que se deducen principios y leyes generales con capacidad predictiva y comprobables experimentalmente”.

### ***Tecnología:***

“1. f. Conjunto de teorías y de técnicas que permiten el aprovechamiento práctico del conocimiento científico”.

“2. f. Tratado de los términos técnicos”.

“3. f. Lenguaje propio de una ciencia o de un arte”.

“4. f. Conjunto de los instrumentos y procedimientos industriales de un determinado sector o producto.

La definición de *ciencia* de este diccionario es precisa y asimilable fácilmente, no así la de *tecnología*, que presenta diferentes formas y puntos de vista. De las que allí se muestran, la más ilustrativa para los propósitos de esta tesis es la combinación de la primera con la cuarta, así se propone la siguiente definición:

***Tecnología:***

*Conjunto de conocimientos científicos y técnicos integrados en equipos, máquinas y dispositivos, orientados a la producción de bienes y servicios útiles para las personas.*

De cualquier manera, la intención de este trabajo no es homologar estas definiciones, sino utilizar los conceptos de forma amplia para explicar y entender los impactos derivados de los avances en el conocimiento científico y tecnológico.

Por otro lado, con la finalidad de establecer criterios homologados a nivel internacional, la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO, 197) definió a las *Actividades Científicas y Tecnológicas* como “*aquellas llevadas a cabo de forma sistemática que están estrechamente relacionadas con la generación, avance, promoción, difusión y aplicación del conocimiento científico y técnico en todos los campos de la ciencia y la tecnología*”, pero los términos *ciencia* y *tecnología* se dan por entendidos [17].

Las actividades científicas y tecnológicas se dividen en tres grandes tipos, de acuerdo con el Manual de Frascati [15]:

1. Investigación y desarrollo experimental (IDE): comprende el trabajo creativo y sistemático llevado a cabo para aumentar el caudal de conocimiento, incluido el conocimiento de la humanidad, la cultura y la sociedad, y el uso de estos conocimientos para crear nuevas aplicaciones.
2. Servicios científicos y tecnológicos (SCYT): son todas las actividades relacionadas con la investigación y el desarrollo tecnológico que contribuyen a la generación, la difusión y la aplicación de los conocimientos científicos y tecnológicos.
3. Educación y enseñanza científica y técnica: Se refiere a todas las actividades de educación y enseñanza de nivel superior: a) no universitario especializado

(estudios técnicos terminales que se imparten después del bachillerato o enseñanza media superior), b) de educación y enseñanza de nivel superior que conduzcan a la obtención de un título universitario (estudios a nivel licenciatura), y c) estudios de postgrado; capacitación y actualización posteriores y de formación permanente y organizada de científicos e ingenieros.

Es pertinente mencionar que en esta tesis no se considera el conocimiento profundo que sólo los científicos e ingenieros poseen de sus respectivos temas de ciencia y tecnología, solamente el conocimiento y entendimiento básico de las personas en torno al tema de ciencia y tecnología, principalmente de las actividades de IDE, así como su opinión resultante sobre sus resultados, aplicaciones y riesgos.

La percepción pública de la ciencia y la tecnología es resultado de la cultura científica y tecnológica de las personas, la cual se adquiere de dos formas básicas, una mediante el conocimiento obtenido de manera formal en la escuela, y la otra a través de la consulta de información realizada por las personas sobre el tema disponible en los medios de comunicación masiva, como son la televisión, radio, internet, periódicos, entre otros. Asimismo, hay otras fuentes de información sobre el tema de ciencia y tecnología que influyen en la percepción, pero que no forman parte de una cultura científica, como lo son las opiniones de personas ajenas al tema, pero que tienen algo que decir al respecto, aun siendo ignorantes del tema.

Una persona que no avanzó mucho en la escuela y que no acostumbra a leer ni escuchar noticias sobre temas de ciencia y tecnología, es muy probable que tenga poca o nula información al respecto y su percepción sobre el tema será ambigua o incluso distorsionada, sobre todo cuando otros factores como la opinión de otras personas no calificadas en el tema (por ejemplo, religiosas) le ofrecen un punto de vista muy particular, e incluso contrario.

En el caso opuesto se encuentran las personas involucradas en temas de ciencia y la tecnología que, además de haber logrado niveles académicos superiores y contar con mayores conocimientos científicos básicos, constantemente están informándose sobre estos temas mediante la consulta en periódicos, revistas y otros medios de información masiva y alcanzan una cultura científica mayor, lo cual redundará en una percepción más precisa sobre el tema, y generalmente más favorable [2].

Cabe mencionar que la percepción pública de la ciencia y tecnología es importante en una democracia participativa, ya que los ciudadanos pueden formar parte de los debates relacionados con las políticas públicas de ciencia y tecnología, para que el gobierno, en cualquiera de sus niveles, apoye desarrollos que pueden ser útiles a la sociedad, el medio ambiente o la economía nacional o local [13].

Asimismo, la participación ciudadana puede frenar el apoyo gubernamental a actividades científicas y tecnológicas en las que se considere que no se está

actuando éticamente, o bien que dañan a los seres vivos o simplemente que no sean útiles para la sociedad [13].

### **1.3 Antecedentes.**

Los primeros esfuerzos por atender la imagen pública de la ciencia y la tecnología surgieron en la década comprendida entre 1950 y 1960, en la cual algunos investigadores sociales y funcionarios públicos del sector ciencia y tecnología de varios países iniciaron trabajos aislados de diseño de instrumentos de medición de la percepción pública de la ciencia y la tecnología, los cuales incluyen de manera relevante la preocupación por medir la alfabetización o el nivel de cultura científica y tecnológica de la sociedad [19].

En ese entonces, los científicos y funcionarios públicos que desarrollaban este tema discutían acerca de cuáles de éstos serían los beneficios de una mayor comprensión pública de la ciencia y la tecnología. De tales debates surgieron argumentos clasificados en cuatro grupos, de acuerdo con lo resumido en el Manual de Antigua [22], que a continuación se transcriben literalmente:

- a. “Razones prácticas: las personas deben tener un buen conocimiento de la ciencia y la tecnología para la vida diaria en sociedades dominadas por el conocimiento científico-tecnológico”.
- b. “Razones culturales: la ciencia es parte de la herencia cultural y ha influenciado profundamente nuestra visión del mundo y del lugar del hombre en él; por ello, se necesita entender qué es la ciencia a fin de comprender la cultura. Además, conocer algo acerca de los objetos y fenómenos del mundo es una fuente de realización personal”.
- c. “Razones económicas: una fuerza de trabajo científica y tecnológicamente alfabetizada es necesaria para una economía basada en el conocimiento”.
- d. “Razones democráticas: las personas necesitan una comprensión de la ciencia y la tecnología para desempeñarse en muchos temas complejos que afectan al ejercicio de la democracia y la ciudadanía en el mundo contemporáneo”.

Los investigadores y funcionarios involucrados en este tema moldearon la medición de la percepción pública hacia la atención de necesidades particulares y nuevos contextos, dando ya sea respuesta, apoyo o bien generando acciones de integración de la percepción con necesidades económicas y sociales, como las agendas de investigación, estrategias metodológicas y estrategias de intervención política identificados con una o más de estas razones.

Con el tiempo, los indicadores estadísticos desarrollados por Estados Unidos a través de la Fundación Nacional para la Ciencia (NSF, por sus siglas en inglés) [14], y por la Unión Europea mediante la aplicación de los Eurobarómetros de ciencia y tecnología [7], se erigieron como la referencia mundial en la medición de la percepción y comprensión pública de este tema. En Estados Unidos los principales indicadores son los relacionados con la comprensión pública de la ciencia (Public Understanding of Science), los cuales se organizan en tres dimensiones de análisis [14]:

- a. Conocimiento.
- b. Interés.
- c. Actitudes.

Cabe mencionar que desde 1972, la NSF incluye un capítulo sobre comprensión y actitudes públicas hacia la ciencia y la tecnología en su informe anual de indicadores [14].

Las encuestas de 1979 en Estados Unidos, así como las encuestas paralelas de 1988 en Reino Unido y Estados Unidos y el Eurobarómetro 38.1 del año 1992 fueron las tempranas y principales referencias en el contexto del desarrollo y posteriores revisiones de la construcción de este tipo de indicadores.

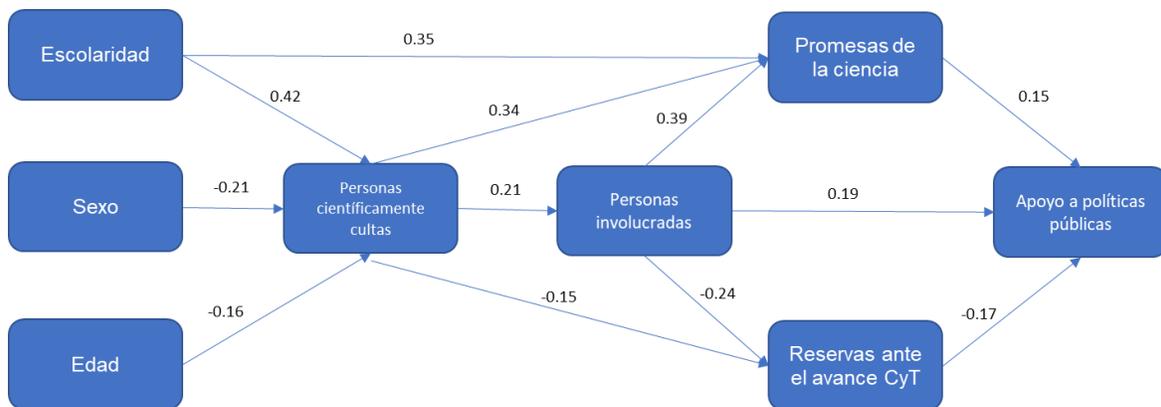
### **1.3.1 Primer estudio de análisis de trayectorias.**

El primer estudio formal sobre la percepción pública de la ciencia y tecnología de alcance internacional lo realizó el profesor John D. Miller [13], de la Universidad del Norte de Illinois, por solicitud de la OCDE, para llevar a cabo un análisis comparativo entre países miembros de dicha organización que contaban con una encuesta del tema.

Dicho estudio utilizó un modelo de participación ciudadana en economías democráticas que permite establecer el perfil de los ciudadanos aptos para participar en debates junto con los tomadores de decisiones, entre los que se encuentran funcionarios públicos, legisladores, magistrados, políticos, universidades y centros de investigación involucrados, así como con líderes de opinión, como científicos, empresarios y periodistas, por ejemplo, para la solución de conflictos derivados de la promulgación de una política pública en el tema de ciencia y tecnología.

Para establecer el perfil ciudadano antes mencionado, Miller desarrolló un modelo estadístico de relaciones causales entre diversas variables llamado *Análisis de trayectorias* (Path analysis), el cual relaciona entre sí a variables sociales (edad, sexo, escolaridad) con factores construidos a partir de variables de conocimiento, consumo de información, percepción pública y apoyo a políticas científicas y tecnológicas.

**Figura 1. Esquema de Análisis de Trayectorias (ejemplo ilustrativo) [13]**



El factor *personas científicamente cultas* se cuantificó mediante la aplicación de una batería de preguntas sobre conocimiento científico básico, así como dos preguntas en las que la persona informante responde a un planteamiento de estudio científico y a un planteamiento probabilístico. Las personas que contestaron acertadamente a todas las preguntas se les asignó una calificación de 100 y a las que respondieron erróneamente a todas las preguntas se les calificó con cero. Las demás calificaciones se distribuyeron mediante “regla de tres”.

Para establecer el factor *personas involucradas*, en primer lugar, se clasificaron a las personas científicamente cultas en tres grandes grupos, las que obtuvieron calificación mayor o igual a 80, las que obtuvieron calificación entre 60 y 79 y las que no alcanzaron 60 de calificación. Así, se incluyeron a las personas del grupo de cultura científica con mejor calificación (mayor o igual a 80) y, adicionalmente se midieron sus hábitos de consumo de información científica y tecnológica, mediante la consulta en revistas, periódicos, programas de radio, televisión, e internet. Lo anterior concluyó que las personas más involucradas son las que, además de tener calificación mayor o igual a 80, consumen más información científica y tecnológica en los diferentes medios.

Los factores *promesas de la ciencia* y *reservas ante el avance científico y tecnológico* son resultado de un proceso de *análisis factorial exploratorio*<sup>1</sup>, en el que se incorporaron las variables de percepción y, como resultado del ejercicio estadístico, se formaron tales factores. El factor *promesas de la ciencia* representa la percepción positiva de las personas en torno al avance científico y tecnológico; por ejemplo, cuando las personas consideran favorable el desarrollo tecnológico

<sup>1</sup> El Análisis factorial es una técnica estadística de reducción de variables usada para explicar las correlaciones entre las variables observadas en términos de un número menor de variables no observadas llamadas factores. Las variables observadas se modelan como combinaciones lineales de factores más expresiones de error [11].

porque facilitará el trabajo, o cuando consideran que el avance del conocimiento científico implicará mejores condiciones de vida. Por su parte, el factor *Reservas ante el avance científico y tecnológico* agrupa a las variables que miden inconformidad o temores de las personas ante el desarrollo de la ciencia y la tecnología como, por ejemplo, la afirmación de que el avance científico y tecnológico desplazará más empleos de los que creará, o bien que no están de acuerdo con las técnicas de experimentación que hacen sufrir a los animales.

### 1.3.2 Estudio de análisis de trayectorias en México.

En 1999, el Consejo Nacional de Ciencia y Tecnología (Conacyt) replicó dicho modelo para el caso de México utilizando los resultados de la *Encuesta Sobre la Percepción Pública de la Ciencia y la Tecnología, ENPECYT 1997* [3]. Cabe mencionar que este estudio fue realizado por la Subdirección de Diseño Estadístico de dicho consejo, a cargo del tesista.

Así, se definieron las variables o factores considerados en el planteamiento teórico de Miller a partir de las preguntas del cuestionario aplicado en la ENPECYT 1997.

Las variables utilizadas en el modelo fueron sexo, edad, escolaridad, cultura científica de la sociedad, público involucrado en la ciencia y la tecnología, promesas de la ciencia, inquietud ante el avance de la ciencia, y opinión sobre política científica.

La construcción de las variables fue determinada en las definiciones antes mencionadas, y utilizando la técnica de análisis de factores. Las escalas consideradas son las siguientes:

**Tabla 2. Clasificación de las variables y sus escalas de medición [3].**

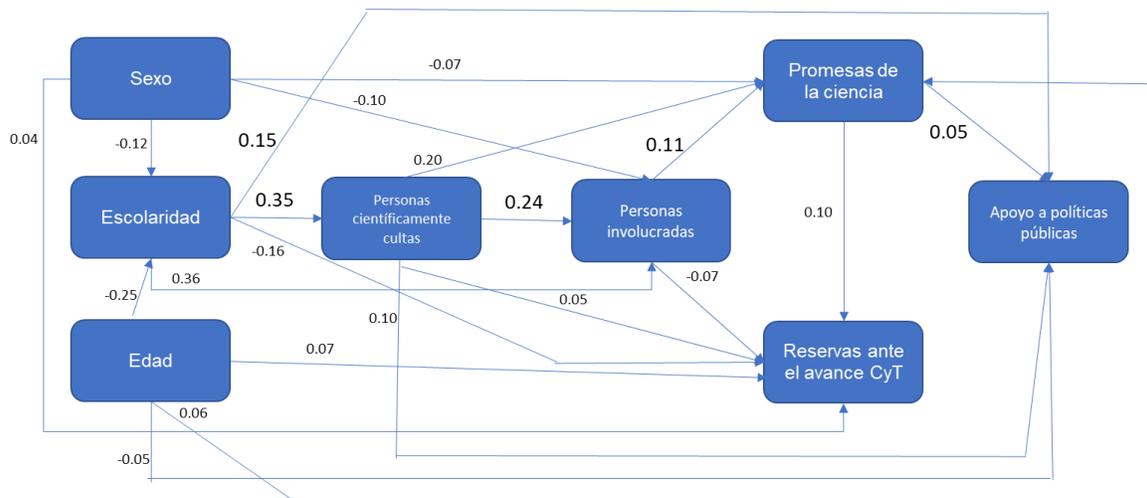
<b>Variable</b>	<b>Escala de medición</b>
Sexo	0 = masculino, 1 = femenino
Edad	años cumplidos
Escolaridad	Número de años cursados en educación formal
Cultura científica	Escala de 0 a 100
Público involucrado	Escala de 0 a 100
Promesas de la ciencia	Escala de 0 a 100
Inquietud ante el avance de la ciencia	Escala de 0 a 100
Opinión sobre política científica	Escala de 0 a 100

En las escalas comprendidas entre 0 y 100, el valor 100 corresponde al nivel más alto alcanzado en cada concepto.

Una vez definidas todas las variables del modelo, se realizó el análisis de trayectorias ya mencionado, con la finalidad de establecer la fuerza de asociación entre las variables, respetando la cadena de causalidad propuesta en el modelo de Miller.

Los principales resultados se muestran de manera resumida en la siguiente figura, en la cual se incorporan únicamente los coeficientes de trayectoria que se obtuvieron con un nivel de significancia de 5%.

**Figura 2. Análisis de Trayectorias de México, 1999 [3].**



El signo de los coeficientes muestra la dirección de la relación. Por ejemplo, de escolaridad a cultura científica el coeficiente es 0.35, lo que indica que las personas con mayor cultura son las que tienen mayor nivel de escolaridad, mientras que de edad a escolaridad se observa el coeficiente  $-0.25$ , lo que significa que la mayor escolaridad la tienen las personas más jóvenes. Asimismo, las personas con cultura científica son más propensas a apoyar a las políticas públicas conforme dicha cultura es mayor, aunque su coeficiente (0.10) es muy bajo.

A partir de este modelo se verifica que una de las trayectorias con mayor peso relativo es la que parte de escolaridad y pasa por cultura científica (0.35), de allí parte a involucramiento en la ciencia (0.24), posteriormente va a promesas de la ciencia (0.11), y termina en política científica y tecnológica (0.05). Cabe mencionar que, conforme avanza la trayectoria, la asociación va siendo cada vez menor, pero se mantiene en relación directa. Así, en términos de los coeficientes de las trayectorias que unen a las variables mencionadas, tanto directa como indirectamente, se puede afirmar que, a mayor escolaridad, mayor cultura científica, mayor involucramiento en la ciencia y actitud más favorable a apoyar las políticas científicas y tecnológicas. Asimismo, si se parte directamente de escolaridad a apoyo a política científica y tecnológica, su coeficiente de trayectoria es 0.15

Los estudios mencionados anteriormente representan avances importantes en la comprensión del fenómeno de la percepción pública de la sociedad a nivel internacional y en México. Permiten cuantificar las relaciones de diferentes factores con el apoyo a las políticas de ciencia y tecnología.

Sin embargo, es pertinente encontrar otras formas complementarias para analizar el tema. Una de estas formas se presenta en esta tesis, cuyo valor adicional consiste en calcular la probabilidad de que una persona con cierto perfil socio económico y de costumbres de consulta de información científica y tecnológica apoye a las políticas en ciencia y tecnología.

## **II FUENTE DE INFORMACIÓN**

En este capítulo se realizará una descripción detallada de la Encuesta sobre la Percepción Pública de la Ciencia y la Tecnología en México, ENPECYT 2013, que es la fuente de datos para llevar a cabo el modelo de regresión logística utilizado en esta tesis. Para tal efecto, se presenta una explicación sobre la importancia de la ENPECYT, su origen, los objetivos que persigue, el diseño del cuestionario y, de forma detallada, el diseño muestral para el cálculo del tamaño de la muestra. Los detalles técnicos son extraídos de forma literal de la síntesis metodológica de dicha encuesta elaborada por el INEGI. Al final del capítulo se presentan algunas observaciones sobre el diseño muestral.

### **2.1 Encuesta sobre la Percepción Pública de la Ciencia y la Tecnología en México, ENPECYT 2013.**

La información estadística generada por las unidades de estado del país debe cumplir con lo dispuesto en el Artículo 3 de la Ley del Sistema Nacional de Información Estadística y Geográfica (SNIEG), que dice: “El Sistema Nacional de Información Estadística y Geográfica, tiene la finalidad de suministrar a la sociedad y al Estado Información de calidad, pertinente, veraz y oportuna, a efecto de coadyuvar al desarrollo nacional” [24]. Para ello, el Conacyt en su calidad de unidad de estado y en cumplimiento con sus compromisos con el SNIEG, desde 1994 realiza de forma sistemática la mayor parte de sus encuestas en colaboración con el INEGI, en particular, la Encuesta Sobre la Percepción Pública de la Ciencia y la Tecnología (ENPECYT), salvo un ejercicio.

Como ya se mencionó en el capítulo anterior, el Conacyt realizó la primera ENPECYT en 1997, en colaboración con una consultora privada. Posteriormente, a partir de 2001 el Consejo ha realizado este ejercicio estadístico de manera sistemática cada dos años en colaboración con el INEGI. Los datos que se utilizan en esta tesis son correspondientes a la ENPECYT 2013. A continuación, se describen los principales componentes de esta encuesta.

#### **2.1.1 Objetivo general de la ENPECYT.**

Recopilar información relevante para la generación de indicadores que midan el conocimiento, entendimiento, consumo de información, percepción y actitud de las personas, relativos a las actividades científicas y tecnológicas.

#### **Objetivos específicos**

Obtener información que permita generar indicadores que midan:

- El conocimiento de las personas sobre conceptos básicos de ciencia y tecnología, así como de planteamientos de procesos científicos y probabilísticos.
- El interés de la sociedad y su entendimiento por la ciencia y la tecnología.
- Actitudes ante el avance científico y tecnológico.
- La disposición de la gente para apoyar el gasto gubernamental en ciencia y tecnología.
- Consulta de medios.
- Conocimiento e imagen del CONACYT.
- El perfil socioeconómico de las personas participantes y su relación con los anteriores indicadores.

### **2.1.2 Cobertura de la encuesta.**

#### **Cobertura sectorial**

El segmento de la población que es denominado “población objetivo” y a la cual se va a encuestar, es: personas de 18 y más años cumplidos, residentes en las viviendas particulares localizadas en áreas urbanas de 100,000 y más habitantes.

#### **Cobertura geográfica**

La encuesta proporciona resultados al agregado de 32 áreas urbanas de interés de 100 000 y más habitantes de las ciudades objeto de estudio de la ENOE.

#### **Cobertura temporal**

Los datos recabados en las viviendas seleccionadas, y que son objeto de estudio de la ENPECYT 2013, corresponden al momento en que se levanta la información en campo.

### **2.1.3 Unidad de observación**

La población objetivo para la aplicación de esta encuesta son las personas de 18 o más años de edad, residentes habituales de las viviendas particulares ubicadas en áreas urbanas de 100,000 y más habitantes, y que además cumplan con el criterio de elegibilidad.

#### **2.1.4 Diseño del cuestionario.**

Para la elaboración del cuestionario se tomaron en cuenta los resultados de las encuestas anteriores, participando en su diseño funcionarios de la Dirección de Análisis Estadístico del Conacyt en conjunto con personal de la Dirección de Estadísticas de Ciencia y Tecnología y de la Dirección de Encuestas Tradicionales del INEGI.

El cuestionario quedó conformado por un total de 74 preguntas distribuidas en las siguientes secciones:

- I. Residentes y hogares.
- II. Datos básicos.
- III. Introducción.
- IV. Comprensión y percepción de la ciencia y la tecnología.
- V. CONACYT.
- VI. Agenda ciudadana de ciencia, tecnología e innovación en México.

#### **2.1.5 Esquema de muestreo.**

El diseño de la muestra para la ENPECYT 2013 se caracteriza por ser probabilístico; en consecuencia, bajo este esquema es factible que los resultados obtenidos de la encuesta se generalizan a toda la población y se puedan medir los errores de las estimaciones obtenidas. A su vez, el diseño es unietápico, estratificado y por conglomerados, donde la unidad última de selección es la vivienda y la unidad de observación es el hogar.

#### **2.1.6 Marco muestral.**

El marco de muestreo utilizado para la ENPECYT 2013 es el Marco Nacional de Viviendas 2012 del INEGI, el cual se construye a partir de información cartográfica y demográfica que se obtuvo del Censo de Población y Vivienda 2010, el cual puede considerarse como una muestra maestra, a partir de la cual se seleccionan las muestras para todas las encuestas en viviendas que realiza el INEGI; las que son denominadas Unidades Primarias de Muestreo (UPM), para después seleccionar en una segunda etapa, las viviendas que integran las muestras de las diferentes encuestas en hogares. Las UPM se conforman por agrupaciones de viviendas con características diferenciadas dependiendo del ámbito al que pertenecen; en el caso del urbano alto se hace de la siguiente manera: El tamaño mínimo de una UPM es de 80 viviendas habitadas y un máximo de 160 y pueden estar formadas por:

- Una manzana

- La unión de dos o más manzanas contiguas de la misma AGEB<sup>2</sup>.
- La unión de dos o más manzanas contiguas de diferentes AGEB de la misma localidad.
- La unión de dos o más manzanas contiguas de diferentes localidades, pero del mismo tamaño de localidad.

### 2.1.7 Estratificación.

Dentro del ámbito urbano alto, se formaron cuatro estratos agrupando todas las UPM del país, considerando las características sociodemográficas de los habitantes de las viviendas, así como sus características físicas y equipamiento, a través de 34 indicadores que fueron construidos con información del Censo de Población y Vivienda 2010, empleando métodos estadísticos multivariados.

### 2.1.8 Tamaño de la muestra.

Para la realización del cálculo del tamaño de muestra, tuvieron que analizarse algunos parámetros que están vinculados con ciertos indicadores de interés, tomándose como referencia la variable “proporción de la población con nivel de escolaridad hasta primaria con interés moderado en el conocimiento de nuevos descubrimientos científicos”, a fin de que las variables con proporciones mayores queden cubiertas de manera automática.

Para el cálculo se utilizó la siguiente expresión:

$$n = \frac{z^2 q \text{ DEFF}}{r^2 p (1 - tnr)}$$

Donde:

n = Tamaño de la muestra.

p = Estimación de la proporción de interés.

q = (1-p), estimación del complemento de p.

z = Valor en tablas estadísticas de la distribución normal para una confianza prefijada.

r = Error relativo máximo esperado.

tnr = Tasa de no respuesta máxima esperada.

DEFF = Efecto de diseño definido como el cociente de la varianza en la estimación del diseño utilizado, entre la varianza obtenida considerando un muestreo aleatorio simple para un mismo tamaño de muestra.

---

<sup>2</sup> AGEB: Área geoestadística básica.

La muestra que se obtuvo fue de 3,159 viviendas y se consideró un nivel de confianza del 90%, una tasa de no respuesta de 15%, un error relativo máximo esperado de 15%, un efecto de diseño de 2.60 y una proporción de 10.43%; ajustándola a 3,200 viviendas distribuidas tal como se observa en la siguiente tabla:

**Tabla 3. Distribución de la muestra de viviendas por área urbana.**

Área urbana	Viviendas	Área urbana	Viviendas
Ciudad de México	100	Tuxtla Gutiérrez	100
Guadalajara	100	Tijuana	100
Monterrey	100	Culiacán	100
Puebla	100	Hermosillo	100
León	100	Durango	100
San Luis Potosí	100	Tepic	100
Mérida	100	Campeche	100
Chihuahua	100	Cuernavaca	100
Tampico	100	Oaxaca	100
Veracruz	100	Zacatecas	100
Acapulco	100	Colima	100
Aguascalientes	100	Querétaro	100
Morelia	100	Tlaxcala	100
Toluca	100	La Paz	100
Saltillo	100	Cancún	100
Villahermosa	100	Pachuca	100

### 2.1.9 Afijación de la muestra.

Esta se realiza en cada una de las ciudades entre los diferentes estratos. La expresión que se utiliza para el cálculo de la afijación de la muestra es el siguiente:

$$n_{ch} = \frac{N_{ch}}{N_c} n_c$$

Donde:

$n_{ch}$  = Número de viviendas en muestra en el h-ésimo estrato, en la c-ésima ciudad.

$n_c$  = Número total de viviendas en muestra, en la c-ésima ciudad.

$N_{ch}$  = Número total de viviendas en el h-ésimo estrato en la c-ésima ciudad.

$N_c$  = Número total de viviendas, en la c-ésima ciudad.

### 2.1.10 Selección de la muestra.

La selección se llevó a cabo de manera independiente para cada área urbana de la siguiente forma:

1. De las  $k_{ch}$  UPM que forman parte de la muestra maestra<sup>3</sup>, se eligieron  $k_{ch}^*$  UPM con igual probabilidad para la ENPECYT 2013.
2. En cada UPM se seleccionaron cinco viviendas con igual probabilidad.
3. En cada una de las viviendas, se seleccionó una persona de 18 años cumplido y más.

a) La probabilidad de seleccionar una vivienda de la  $i$ -ésima UPM, del  $h$ -ésimo estrato de la  $c$ -ésima ciudad es:

$$P\{V_{chi}\} = \frac{k_{ch} m_{chi}}{m_{ch}} \frac{k_{ch}^*}{k_{ch}} \frac{5}{m_{chi}^*} = \frac{5k_{ch}^* m_{chi}}{m_{ch} m_{chi}^*}$$

b) La probabilidad de seleccionar una persona de 18 años cumplidos y más de la  $j$ -ésima vivienda, de la  $i$ -ésima UPM, del  $h$ -ésimo estrato, de la  $c$ -ésima ciudad,  $H_{chij}$  es:

$$P\{H_{chij}\} = \frac{k_{ch} m_{chi}}{m_{ch}} \frac{k_{ch}^*}{k_{ch}} \frac{5}{m_{chi}^*} \frac{1}{H} = \frac{5k_{ch}^* m_{chi}}{m_{ch} m_{chi}^* H}$$

### 2.1.11 Cálculo de los factores de expansión.

El factor de expansión<sup>4</sup> está dado por:

$$F_{chij} = \frac{H m_{ch} m_{chi}^*}{5k_{ch}^* m_{chi}}$$

Donde:

- $k_{ch}$  = Número de UPM seleccionadas para el marco de la muestra maestra en el  $h$ -ésimo estrato en la  $c$ -ésima ciudad.
- $m_{chi}$  = Número de viviendas en la  $i$ -ésima UPM, en el  $h$ -ésimo estrato, en la  $c$ -ésima ciudad.
- $m_{ch}^*$  = Total de viviendas en el  $h$ -ésimo estrato en la  $c$ -ésima ciudad.
- $m_{chi}^*$  = Total de viviendas en la  $i$ -ésima UPM, en el  $h$ -ésimo estrato en la  $c$ -ésima ciudad al momento del levantamiento de la ENPECYT 2013.

<sup>3</sup> Del total de UPM que integran el marco de propósitos múltiples se seleccionaron con probabilidad proporcional a su tamaño  $k_{ch}$  UPM para localidades de 100 000 y más habitantes.

<sup>4</sup> El factor de expansión es el inverso de la probabilidad de selección.

- $k^*_{ch}$  = Número de UPM seleccionadas en el h-ésimo estrato en la c-ésima ciudad.
- $H$  = Número de personas de 18 años cumplidos y más en la vivienda seleccionada.

### 2.1.12 Ajustes a los factores de expansión.

Los factores de expansión se ajustan para los siguientes conceptos de acuerdo al procedimiento anterior:

**Ajuste por no respuesta a nivel vivienda:** este ajuste atribuido al informante se hace a nivel UPM en cada dominio y se utiliza la siguiente expresión:

$$F'_{chi} = F_{chi} \frac{V_{chi}}{V^*_{chi}}$$

Donde:

- $F'_{chi}$  = Factor de expansión corregido por no respuesta para las viviendas de la i-ésima UPM, del h-ésimo estrato en la c-ésima ciudad.
- $F_{chi}$  = Factor de expansión de la i-ésima UPM, del h-ésimo estrato, en la c-ésima ciudad.
- $V_{chi}$  = Número de viviendas habitadas seleccionadas en la i-ésima UPM, en el h-ésimo estrato, en la c-ésima ciudad.
- $V^*_{chi}$  = Número de viviendas habitadas seleccionadas en la i-ésima UPM, en el h-ésimo estrato, en la c-ésima ciudad para las cuales se obtuvo respuesta.

**Ajuste por no respuesta a nivel persona:** estos fueron corregidos a nivel UPM, ya que algunas de las personas de 18 años cumplidos y más seleccionadas no contestaron. La expresión utilizada es:

$$F''_{chij} = F'_{chij} \frac{H}{H^*}$$

Donde:

- $F''_{chij}$  = Factor de expansión corregido por no respuesta a nivel personas, de la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato en la c-ésima ciudad.
- $F'_{chij}$  = Factor de expansión a nivel persona seleccionada en la UPM, de la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, en la c-ésima ciudad.
- $H$  = Número de personas de 18 años cumplidos y más, seleccionadas por UPM.
- $H^*$  = Número de personas de 18 años cumplidos y más, seleccionadas por UPM, para las que se obtuvo respuesta.

*Ajuste por proyección:* para asegurar que se obtendrá una población determinada por la proyección, se corrigen los factores de expansión ajustados al punto medio del levantamiento a través de la siguiente expresión:

$$F_D''' = F_D'' \frac{PROY_D}{PEXP_D}$$

Donde:

- $F_D'''$  = Factor de expansión corregido por proyección en el dominio D.
- $F_D''$  = Factor de expansión corregido por no res- puesta en el dominio D.
- $PROY_D$  = Población en el dominio D, según proyección.
- $PEXP_D$  = Población total a la que expande la encuesta en el dominio D.

### 2.1.13 Estimadores.

El estimador del total de la característica X es:

$$\hat{X} = \sum_c \sum_h \sum_i \sum_j F_{chij} X_{chij}$$

Donde:

- $F_{chij}$  = Factor de expansión final de la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, en la c-ésima ciudad.
- $X_{chij}$  = Valor observado de la característica de interés X de personas seleccionadas, en la j-ésima vivienda, de la i-ésima UPM, del h-ésimo estrato, en la c-ésima ciudad.

El siguiente estimador de razón se utiliza para la estimación de proporciones, tasas y promedios:

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

Donde  $\hat{Y}$  se define en forma análoga a  $\hat{X}$ .

### 2.1.14 Estimaciones de errores de muestreo.

El método de conglomerados últimos se utilizó para la evaluar los errores de muestreo de las principales estimaciones nacionales, el cual se basa en que la mayor contribución a la varianza de un estimador en un diseño bietápico, es la que se presenta entre las UPM. El término “conglomerados últimos” se refiere al total de unidades en muestra de una UPM.

Las precisiones de los estimadores de razón se obtienen aplicando el método de conglomerados últimos junto con el método de series de Taylor mediante la siguiente expresión:

$$\hat{V}(\hat{R}) = \frac{1}{\hat{V}^2} \sum_c \left\{ \sum_h \frac{k_{ch}}{k_{ch}-1} \sum_i^{k_{ch}} \left[ \left( \hat{X}_{chi} - \frac{1}{k_{ch}} \hat{X}_{ch} \right) - \hat{R} \left( \hat{Y}_{chi} - \frac{1}{k_{ch}} \hat{Y}_{ch} \right) \right]^2 \right\}$$

Donde:

- $\hat{X}_{chi}$  = Total ponderado de la variable de estudio X en la i-ésima UPM, en el h-ésimo estrato, en la c-ésima ciudad.
- $\hat{X}_{ch}$  = Total ponderado de la variable de estudio X en el h-ésimo estrato, en la c-ésima ciudad.
- $k_{ch}$  = Número de UPM en el h-ésimo estrato, en la c-ésima.

Estas definiciones son análogas para la variable de estudio Y. La expresión utilizada para el cálculo de la varianza del estimador de un total es la siguiente:

$$\bar{V}(\bar{X}_{NAL}) = \sum_c \sum_{h=1}^{L_c} \frac{k_{ch}}{k_{ch}-1} \sum_{i=1}^{k_{ch}} \left( \bar{X}_{chi} - \frac{1}{k_{ch}} \bar{X}_{ch} \right)^2$$

Las estimaciones de la desviación estándar (D.E.), coeficiente de variación (C.V.) y efecto de diseño (D.E.F.F.) se calculan mediante las siguientes expresiones:

$$DE = \sqrt{\hat{V}(\hat{\theta})} \quad CV = \frac{\sqrt{\hat{V}(\hat{\theta})}}{\hat{\theta}}$$

$$DEFF = \frac{\hat{V}(\hat{\theta})}{\hat{V}(\hat{\theta})_{MAS}}$$

Donde:

- $\hat{\theta}$  = Estimador del parámetro poblacional.
- $\hat{V}(\hat{\theta}_{MAS})$  = Estimador de la varianza bajo muestreo aleatorio simple.

Finalmente, el intervalo de confianza de 90% se calcula mediante la expresión:

$$I_{1-\alpha} = \left( \hat{\theta} - 1.645 \sqrt{\hat{V}(\hat{\theta})}, \hat{\theta} + 1.645 \sqrt{\hat{V}(\hat{\theta})} \right)$$

Donde el nivel de significancia es de  $\alpha = 0.102$ .

## 2.2 Observaciones al diseño muestral.

Desde 2001, el INEGI ha colaborado con el CONACYT para el levantamiento de la ENPECYT, realizando este ejercicio cada dos años. De esa manera, hasta el año 2013 se habían llevado a cabo siete encuestas, las cuales tienen en común tres características recurrentes en su diseño muestral que es conveniente observar. La primera es que hay consistencia en la presentación de su metodología de diseño, siendo exactamente la misma, sin variaciones. Otra es que las zonas urbanas seleccionadas también han sido las mismas y, finalmente, la tasa de no respuesta siempre ha sido 15 por ciento.

En relación con el diseño muestral, es un acierto metodológico que se conserve en todas las encuestas, pues homologa los resultados y permite poder realizar comparaciones temporales entre cada uno de los ejercicios estadísticos.

Sin embargo, dada la naturaleza aleatoria de los cálculos entre cada ejercicio, la selección de zonas urbanas debería ser variable.

Asimismo, la experiencia observada entre las encuestas con información disponible debe resultar en un ajuste de la tasa de no respuesta, ya que se ha observado que varía entre 6.7 y 9.6 por ciento, valores alejados del 15 por ciento utilizado recurrentemente. En particular, la tasa de no respuesta observada de la ENPECYT 2013 [5] fue de 9.6 por ciento, como se puede apreciar en la siguiente tabla.

**Tabla 4. Distribución de la tasa de no respuesta observada en los ejercicios de la ENPECYT [5].**

Año	N calculada	N levantada	Tasa de no respuesta
2001	nd	2,552	
2003	nd	2,727	
2005	nd	2,856	
2007	nd	2,943	
2009	3,159	2,935	7.1
2011	3,159	2,947	6.7
2013	3,159	2,857	9.6

Otro argumento a favor de ajustar la tasa de no respuesta es la evolución de la muestra observada, ya que de 2001 a 2007 reportó crecimiento promedio de 4.9 por ciento anual, al pasar de  $n=2,552$  en 2001 a  $n=2,943$  en 2007, para estabilizarse en

2009 y 2011, y luego experimentar una leve caída en 2013. Sin embargo, dicho descenso no es para considerar una tasa de no respuesta tan alta como la utilizada por el INEGI.

### III DESARROLLO DEL TEMA

En el presente capítulo se presentarán algunos modelos orientados a entender a la participación ciudadana, en particular en lo referente a su apoyo a las políticas públicas en ciencia y tecnología, con base en diversos supuestos conceptuales. En especial se relata el modelo del Manual de Antigua [22] y el modelo alternativo elaborado por el tesista. Ambos modelos junto con el enfoque de Miller [13] dan como resultado el modelo conceptual y estadístico utilizado en este trabajo, el cual se mostrará detallado, enfatizando en la construcción de la variable dependiente *Apoyo a políticas públicas en ciencia y tecnología*.

#### 3.1 Generalidades.

Esta tesis utilizará variables de conocimiento, comprensión, consulta de información y percepción pública de ciencia y tecnología que inciden en el apoyo a políticas públicas por parte de la ciudadanía en general.

En este contexto, de acuerdo con el Manual de Antigua, *“La ciudadanía ha cobrado protagonismo creciente como categoría política, se moviliza, discute, reclama e interviene en la arena pública de distintas maneras. A consecuencia de ello, los gobiernos, pero también las industrias y otras instituciones públicas y privadas, intentan incluirla en la gestión política generando mecanismos de participación y consulta. No es extraño que ello ocurra si se tienen en cuenta las fuertes presiones sociales que se ejercen para abrir los procesos de tomas de decisión más allá de la esfera de los expertos y los poderes establecidos”* [22]. De esta manera, la participación ciudadana en la discusión de políticas públicas de ciencia y tecnología requiere que las personas tengan altos niveles de información en tópicos de ciencia y tecnología, en particular conocimientos básicos y comprensión de planteamientos de experimentos científicos y de eventos probabilísticos.

Por su parte, Miller [13] estableció que el nivel alto de conocimiento científico básico y la comprensión científica y probabilística son necesarias para la alfabetización científica, de tal manera que aquella persona que cuente con este nivel será capaz de leer y comprender la sección de ciencia *Martes* del New York Times.

#### 3.2 Modelo de cultura científica del Manual de Antigua.

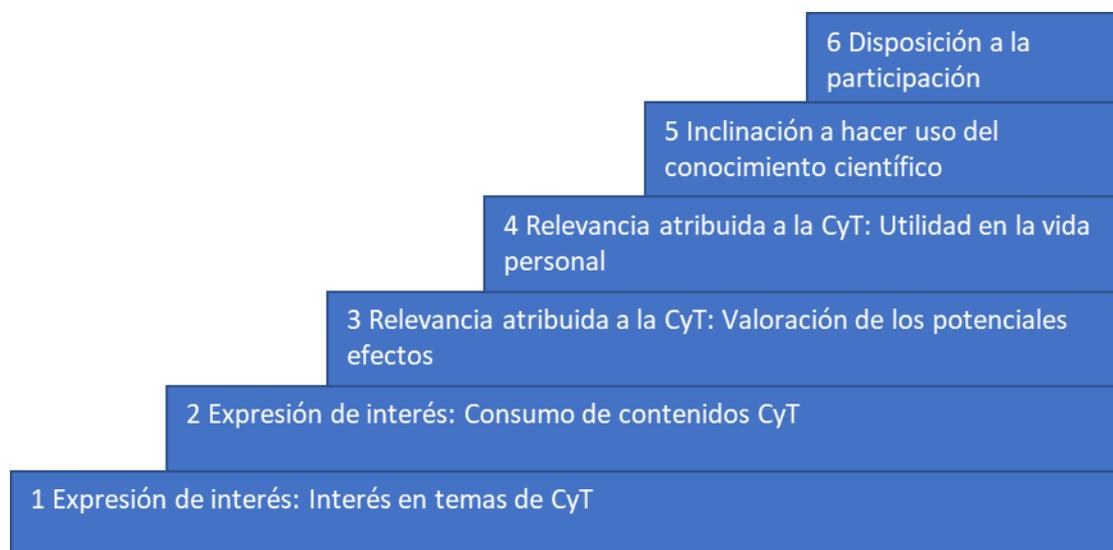
A diferencia de los estudios de análisis de trayectorias de Miller [13] y del Conacyt [3] antes mencionados, en el presente trabajo se desarrolla un modelo de percepción pública y cultura científica que incorpora algunos elementos de Miller y otros del Manual de Antigua, el cual representa la propuesta iberoamericana de medición de estos temas y es el único manual a nivel mundial dedicado a la medición de la percepción pública de la ciencia y la tecnología. Sin embargo, esta

tesis realizará una crítica a la teoría de dicho manual y establecerá su propio criterio, lo cual se presentará más adelante.

En principio, el Manual de Antigua define a la *Cultura Científica* como “*un fenómeno multidimensional resultante de un proceso de apropiación social de la ciencia*” [22]. Sin embargo, no presenta una definición concreta del término *Apropiación Social*. En este sentido, no es necesario usar dicho término pues, aunque estuviera bien definido, en esta tesis se asume que la cultura científica es resultado de la formación académica y de la consulta recurrente de información en los medios de comunicación masiva, lo cual implica conocimientos y opiniones sobre el tema de ciencia y tecnología, siendo un proceso no estático sino dinámico en el tiempo.

Este manual brinda un esquema de la cultura científica y tecnológica mediante una modelo de *escalera*, en la que cada escalón sucede al previo de forma lineal, iniciando con la expresión de interés en temas científicos y tecnológicos, continuando con la consulta de información científica, posteriormente valorando los efectos potenciales del avance científico y tecnológico, lo cual da paso a las valoraciones de la ciencia y la tecnología en la utilidad en la vida personal, continuando con la inclinación hacia el uso del conocimiento científico, y finalizando con la disposición de las personas a la participación ciudadana en estos temas [22].

**Figura 3. Modelo de Escalera de Cultura Científica del Manual de Antigua, 2012. [22]**



Según el Manual de Antigua, “*Los niveles reflejan un cierto ordenamiento gradual de asimilación significativa o enculturación científica y no un proceso genético de desarrollo individual, debido a la interacción múltiple entre ellos dado que el interés influye en el consumo, éste en la importancia atribuida, que a su vez hace aumentar el interés, etcétera*” [22].

### **3.3 Crítica al modelo del Manual de Antigua.**

Según lo expresado en el párrafo anterior, se deduce que hay algunos aspectos que el Manual de Antigua no toma en cuenta. El primero es que no puede existir interés en algo que no se conoce. Para despertar el interés en ciencia y tecnología, las personas deben haber sido informadas sobre el tema, en primer lugar, en la escuela y, posteriormente, en los medios de comunicación masiva. De esta manera se observa que el manual no está considerando a la educación formal como un factor de cultura científica.

Asimismo, no es conveniente considerar que la cultura científica deba ser un proceso escalonado. En particular, la disposición a la participación ciudadana en debates, propuestas, elaboración, votación y seguimiento de políticas públicas en ciencia y tecnología es un factor que depende no necesariamente del escalonamiento de los factores antes mencionados, ya que cabe la posibilidad de “saltarse” uno o dos escalones sin demérito de su opinión.

Otro aspecto no considerado es el de la temporalidad. Una persona tendrá una posición frente a la ciencia y la tecnología igual o diferente en dos momentos del tiempo, dependiendo del consumo de información científica y tecnológica que realice en ese intervalo de tiempo. Es decir, la cultura científica y tecnológica es un proceso cambiante, como lo es el mismo conocimiento científico, el desarrollo tecnológico y muchos otros aspectos de la vida cotidiana, como las estructuras sociales, los modelos económicos, etcétera.

Este manual no da trato individual a las variables, sino que las agrupa en factores y cada factor depende del anterior, como ya se mencionó. Así, la participación ciudadana está en función del factor inmediato anterior y éste del previo, y así sucesivamente.

### **3.4 Modelo alternativo de cultura científica.**

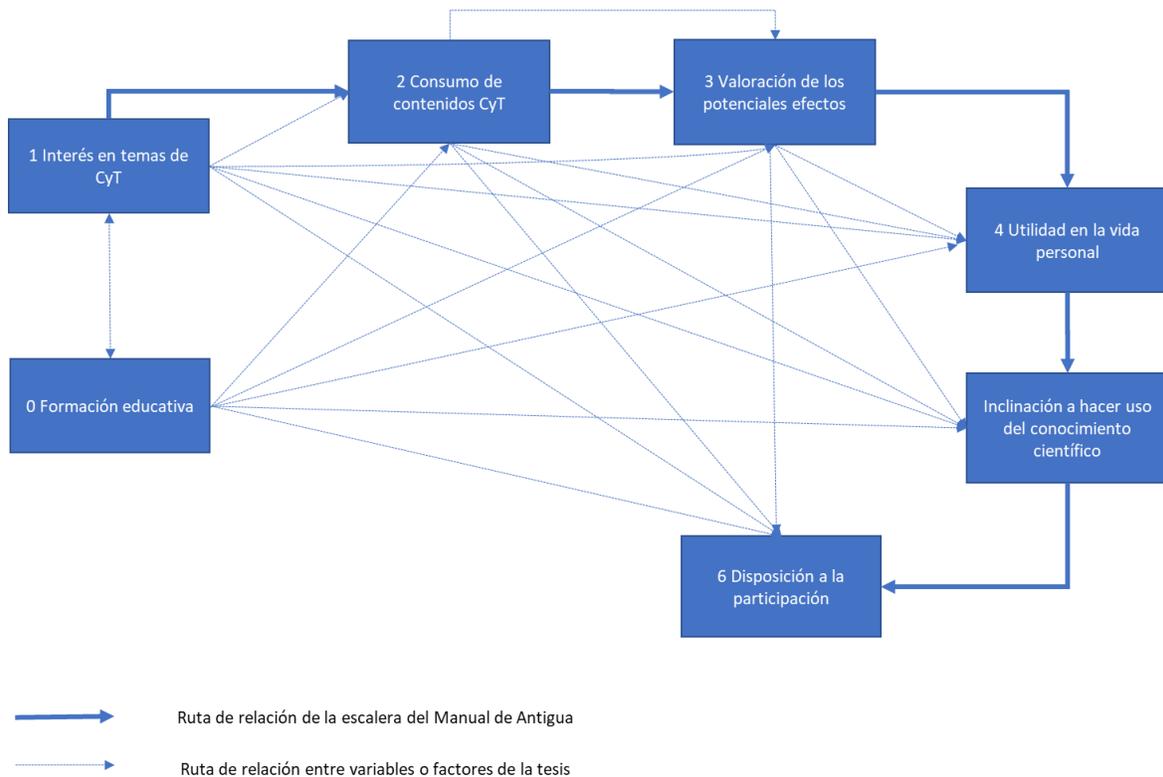
A diferencia de lo que establece el Manual de Antigua, en esta tesis se presenta un modelo alternativo, el cual relaciona a todas las variables y factores seleccionados en forma de red, no lineal.

Este modelo considera que todos los factores y todas las variables están relacionadas entre sí de alguna manera, ya sea en un solo sentido o de forma biunívoca.

Además, incluye a la variable *formación educativa* como parte de la cultura científica, la cual no es considerada por el Manual de Antigua, como se mencionó con anterioridad.

En la Figura 4 se muestran flechas con mayor grosor, que representan a la escalera del Manual de Antigua, mientras que las flechas con menor grosor son las relacionadas con el modelo de red, que es muy parecido al modelo de Miller.

**Figura 4. Cultura científica, Esquema alternativo [22]**



### 3.5 Modelo conceptual y estadístico.

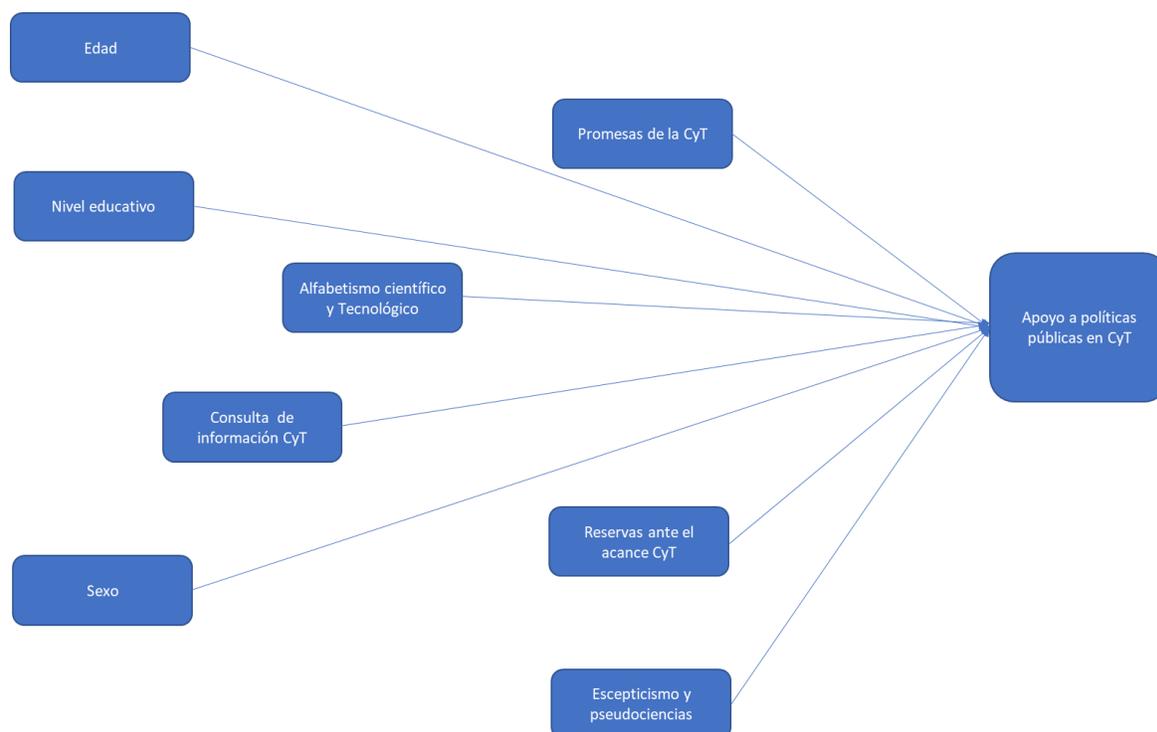
En el capítulo I se mencionaron como antecedentes a los modelos de análisis de trayectorias realizados por Miller y por el Conacyt, y al realizar convenientemente la combinación de su estructura conceptual con la del Manual de Antigua, propician la propuesta conceptual de la presente tesis.

Debido a que el análisis de trayectorias ya fue realizado por el Conacyt, el presente modelo no recurrirá a tal técnica, sino a al uso de la regresión logística para el análisis del modelo. De esta manera, es importante tener en mente que esta tesis busca identificar los factores que incidan directamente (positiva o negativamente) en el apoyo a políticas públicas en ciencia y tecnología, y sobre todo tener la oportunidad de medir la probabilidad de dicho apoyo, dependiendo del tipo de persona que se analice.

En otras palabras, la **Regresión logística multivariada**, también conocida como **Regresión logística** o **Modelo logit**, permite hacer la estimación de los parámetros que miden la incidencia de diferentes variables en el apoyo a las políticas públicas, así como la medición puntual de la probabilidad de apoyo a las políticas públicas en ciencia y tecnología.

Dada la disponibilidad de datos y atendiendo al modelo estadístico propuesto, se presenta el diagrama a analizar en la presente tesis.

**Figura 5. Diagrama del Modelo conceptual y estadístico.**



Como se aprecia en la figura anterior, el apoyo a las políticas públicas en ciencia y tecnología es una variable que depende a su vez de cada una de las variables consideradas. Las variables de percepción pública se agruparon, en términos ilustrativos, en grupos o factores como los que resultarían de un ejercicio de análisis factorial. Las variables agrupadas en *promesas de la ciencia* son las número 9, 12, 13, 14 15, 16, 17 y 20. Las correspondientes a *reservas ante el avance CyT* son la 10, 11, 18, 21 y 23; y las agrupadas en *escepticismo y pseudociencias* son la 19, 22, 24 y 25.

### 3.5.1 Variables del Modelo conceptual y estadístico.

Las variables consideradas para el análisis del modelo son:

#### **Variable dependiente:**

Apoyo a políticas públicas en ciencia y tecnología.

#### **Variables independientes:**

##### *Sociodemográficas:*

1. Sexo: toma los valores 0=mujer y 1=hombre.
2. Formación educativa: número de años cursados en la escuela.
3. Edad: años cumplidos.

##### *Consulta de información científica y tecnológica en medios de comunicación masiva:*

4. Consulta de CyT en televisión: (porcentaje de tiempo de consulta de CyT en relación con la consulta semanal de televisión).
5. Consulta de CyT en radio: porcentaje de tiempo de consulta de CyT en relación con la consulta semanal de radio.
6. Consulta de CyT en periódicos: porcentaje de tiempo de consulta de CyT en relación con la consulta semanal de periódicos.
7. Consulta de CyT en revistas: porcentaje de tiempo de consulta de CyT en relación con la consulta semanal de revistas.

##### *Cultura científica o alfabetismo científico:*

8. Calificación de cultura científica o alfabetismo científico: escala de calificación de 0 a 100 sobre conocimientos básicos de ciencia y planteamientos científicos y probabilísticos.

##### *Percepción pública:* Medida con base en la evaluación del grado de acuerdo con las afirmaciones siguientes

(Muy de acuerdo = 4, De acuerdo = 3, En desacuerdo= 2 Muy en desacuerdo = 1, No sabe = 0)

9. Gracias a la ciencia y la tecnología habrá más oportunidades para las próximas generaciones.
10. Los descubrimientos tecnológicos tarde o temprano destruirán el planeta.
11. Debido a sus conocimientos, los investigadores científicos tienen un poder que los hacen peligrosos.
12. Los científicos deben ser libres de llevar a cabo sus investigaciones a su antojo, siempre y cuando las hagan bajo reglas éticas.

13. En México debería haber más gente trabajando en investigación y desarrollo tecnológico.
14. Los científicos y los empresarios deberían cooperar más entre sí.
15. Los científicos deberían interesarse más en patentar sus investigaciones y en el uso que se les dé.
16. Debería haber más mujeres dedicadas a la investigación científica en nuestro país.
17. El gobierno debe impulsar que las personas participen en debates sobre la asignación de presupuesto para ciencia y tecnología.
18. La ciencia tiene una imagen muy negativa para la sociedad.
19. Confiamos demasiado en la fe y muy poco en la ciencia.
20. La aplicación de la ciencia hace que nuestro modo de vida cambie demasiado rápido.
21. Algunos números son de la suerte.
22. El desarrollo tecnológico origina una manera de vivir artificial y deshumanizada.
23. Algunos de los objetos voladores no identificados que se han reportado, son en realidad vehículos espaciales de otras civilizaciones.
24. Algunas personas poseen poderes psíquicos.
25. Existen medios adecuados para el tratamiento de enfermedades que la ciencia no reconoce (acupuntura, quiropráctica, homeopatía y limpias).

Es importante detallar la construcción de la variable dependiente *apoyo a políticas públicas en ciencia y tecnología*, ya que proviene de dos preguntas, las cuales fueron integradas en una sola, que además fue dicotomizada por conveniencia analítica.

La variable dependiente *Y* consiste en la opinión de las personas en torno al monto de gasto que realiza el gobierno para apoyar tanto a la investigación científica como al desarrollo de tecnología propia. Se asume que cuando manifiestan que es Muy poco dicho monto, es porque consideran que debería ser mayor; es decir, apoyan a las políticas públicas en ese sentido.

Lo anterior proviene de la Pregunta 30, Sección E del cuestionario de la ENPECYT 2013, considerando a los reactivos 4 y 8, cuyo texto e instrucciones se mencionan a continuación.

Voy a nombrar algunos gastos que hace el gobierno, para cada uno de ellos, dígame si cree que se está asignando o invirtiendo muy poco dinero, el monto correcto o demasiado dinero.

**Instrucciones:** Anote el código según corresponda

Muy poco = 1, Monto correcto = 2, Demasiado = 3, No sabe = 4

- 1 Satélites de comunicación .....
- 2 Reducción de la contaminación .....
- 3 Mejoras en los servicios de salud .....
- 4 Apoyos a la investigación científica .....**
- 5 Mejoras en el sistema educativo .....
- 6 Dotar a la población de acceso universal a las tecnologías de la información (computadoras, líneas telefónicas, servicios vía Internet, etc.) .....
- 7 Reducción de la pobreza y atención a los adultos mayores.....
- 8 Apoyos a las empresas para que desarrollen tecnología propia....**

Así, las personas que respondieron la opción Muy poco=1, ya sea en el reactivo 4 *Apoyos a la investigación científica* o en 8 *Apoyos a las empresas para que desarrollen tecnología propia*, se considera como apoyo a las políticas públicas de ciencia y tecnología. A las personas que respondieron cualquier otra cosa se les considera como que no apoyan tales políticas.

Para ello, se consideraron a cada ítem como una variable que se codificó con 1 para el caso de apoyo a las políticas y 0 al no apoyo y posteriormente se construyó la variable dependiente que captura la unión de dichas variables con la misma codificación, de tal manera que la variable dependiente  $Y$  es dicotómica.

De lo anterior, la variable dependiente  $Y$  es:

$$Y = Y_C \cup Y_T$$

Donde:

$Y$  = apoyo a políticas públicas en ciencia y tecnología.

$Y_C$  = apoyo a políticas en ciencia.

$Y_T$  = apoyo a políticas en tecnología.

Las tres variables toman los valores 0 para no apoyo y 1 para apoyo de sus respectivas políticas.

De esta manera, la construcción de la variable dependiente  $Y$  cumple con las condiciones del siguiente cuadro.

**Tabla 5. Construcción de la variable dependiente mediante lógica**

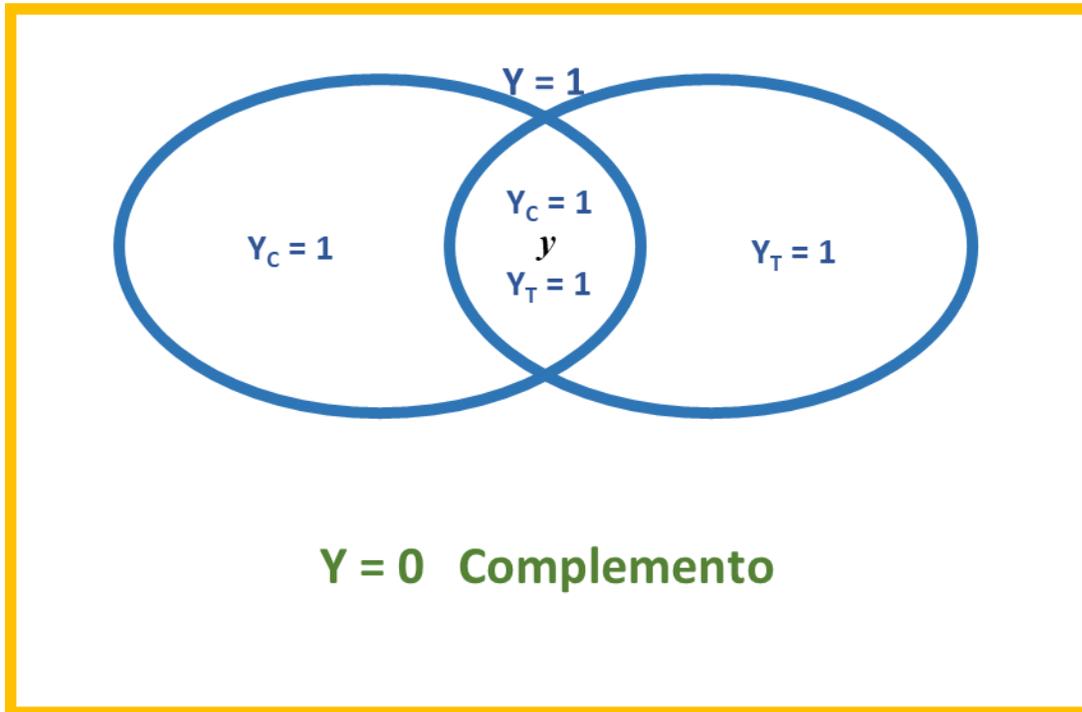
	$Y_C$	$Y_T$		$Y$
<i>Si</i>	1	1	<i>Entonces</i>	1
	1	0		1
	0	1		1
	0	0		0
	0	0		0

O mediante un esquema de conjuntos:

$$Y = Y_C \cup Y_T$$

Para  $Y=1$  y  $Y=0$

**Figura 6. Construcción de la variable dependiente mediante conjuntos**



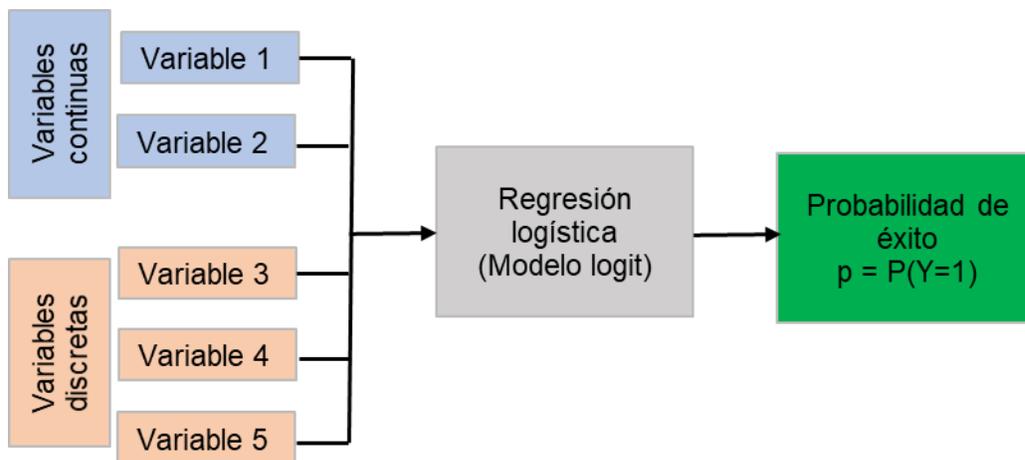
## IV REGRESIÓN LOGÍSTICA MULTIVARIANTE (MODELO LOGIT)

En el presente capítulo se realizará una exposición del modelo logístico con el cuál se llevaron a cabo las estimaciones de parámetros útiles para explicar el modelo teórico empleado en este trabajo. Para ello, se compara su conveniencia con la de otros modelos similares y su pertinencia. Posteriormente se presentará el modelo estadístico con su método de estimación, la interpretación de los coeficientes y sus ajustes correspondientes.

### 4.1 Introducción.

La Regresión Logística Multivariante forma parte de la familia de modelos de regresión lineal, con la particularidad de que su variable dependiente es binaria. Es decir, los resultados que se obtienen son la medición de la probabilidad de un éxito o, complementariamente, de un fracaso [11].

**Figura 7 Esquema del Modelo de Regresión Logística Multivariante.**



De acuerdo con la figura anterior,  $p=P(Y=1)$  es la probabilidad de éxito, donde  $Y$  es la variable dependiente binaria y toma los valores  $Y=0$  si es fracaso y  $Y=1$  si es éxito.

Cabe mencionar que las variables independientes pueden ser cualitativas, cuantitativas, o bien de ambos tipos. En caso de variables categóricas que pueden adoptar varios valores excluyentes, por ejemplo, nivel económico: alto, medio y bajo; sus alternativas se pueden modelar al definir dos variables 0,1. La variable Nivel alto adoptará el valor 1 si está presente, lo cual implica que automáticamente se excluye el nivel medio y el bajo. Si la variable Nivel alto adopta el valor 0 entonces es factible que la variable Nivel medio adopte el valor 1 y queda excluida también la posibilidad de Nivel bajo. En caso de que tanto Nivel alto, como Nivel medio sean 0, significará que automáticamente se trata de un caso de Nivel bajo. Así, no se requiere definir específicamente una variable Nivel bajo, pues ésta queda definida por la ausencia de las otras dos clases. En general, para modelar una variable

categorica de  $K$  valores excluyentes, se requieren  $K-1$  variables indicadoras [23]. Esquematizando el ejemplo mencionado en este párrafo:

**Tabla 5. Construcción de dos variables binarias a partir de una variable categorica de 3 valores excluyentes [23].**

	Nivel alto	Nivel Medio	Nivel Bajo
Nivel alto	1	0	0
Nivel Medio	0	1	0
Nivel Bajo	0	0	1

Para efectos prácticos, cabe mencionar que el paquete estadístico SPSS, el cual fue utilizado en esta tesis, considera la presencia de variables categoricas cuyos valores se identifican por abreviaturas alfanuméricas o por etiquetas numéricas, en cuyo caso hay que especificar al paquete que se trata de variables categoricas. Al momento de indicar a SPSS que se incluyen variables categoricas se generan automáticamente  $K-1$  variables indicadoras que tendrán presencia en el modelo.

De acuerdo con Daniel Peña [18] la Regresión Logística Multivariante tiene tres objetivos básicos:

1. Obtener una estimación no sesgada o ajustada de la relación entre la variable dependiente (o resultado) y una o más variables independientes.
2. Evaluar varios factores simultáneamente que estén presumiblemente relacionados de alguna manera (o no) con la variable dependiente, y conocer su papel y su efecto de forma ajustada.
3. Construir un modelo y obtener una ecuación con fines de predicción o cálculo del riesgo, de manera que éste pueda estimarse para un nuevo individuo con una cierta validez y precisión.

## **4.2 Modelo de Regresión Logística en relación con el Modelo de Regresión Lineal Múltiple.**

Antes de definir formalmente al Modelo de Regresión Logística (Modelo Logit), diversos autores afirman que este modelo [11] [18] [23] forma parte de la familia del Modelo de Regresión Lineal Múltiple (MRLM), el cual asigna a una variable independiente  $Y$  un algoritmo lineal en los parámetros, en relación con las variables independientes. Por ejemplo:

$$Y = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_k X_k + \varepsilon$$

O bien,

$$Y = \alpha_0 + \alpha_1 X_1 + \alpha_{23} X_2 X_3 + \alpha_4 X_4 + \varepsilon$$

Donde:

$Y$  es la variable dependiente.

$X_i$  es la  $i$ -ésima variable independiente.

$\alpha_i$  es el  $i$ -ésimo parámetro a estimar.

$\varepsilon$  es el error aleatorio o residuo.

El MRLM requiere el cumplimiento de varias hipótesis para que sus estimaciones sean aceptadas como estadísticamente significativas y, por lo tanto, explicativas. Entre tales hipótesis se encuentran la normalidad y la homocedasticidad de los residuos. Es decir, que los errores se comporten bajo la distribución normal y que puedan estandarizarse de manera que su media sea cero y su desviación estándar sea 1.

Para cumplir con lo anterior, las variables independientes deben ser continuas, o a lo más discretas numerables, de preferencia con escala amplia; por ejemplo, la edad de la población de un país, cuyo rango de medición va desde 0 años hasta 99 o más. Así, el uso de variables independientes discretas del tipo nominal no es aceptado en este modelo.

Sin embargo, el Modelo Logit asume heteroscedasticidad (varianza variable) y el uso de variables independientes discretas, sean nominales u ordinales, así como de variables continuas. En otras palabras, las variables independientes no necesariamente tienen una distribución conjunta normal multivariante [11] [18] [23].

Por otro lado, el MRLM puede asumir a la variable dependiente como dicotómica, pero presenta un problema de medición una vez que se ha realizado una estimación de los parámetros, ya que, al sustituir los datos de un registro a evaluar en la ecuación resultante, difícilmente se reportarán valores 0 o 1 en la variable dependiente. De hecho, es frecuente que los resultados sean en el rango de los números reales.

Por lo tanto, en esta tesis se utilizará el Modelo Logit debido a la naturaleza de la mayoría de las variables seleccionadas (ordinales), y porque se busca calcular la probabilidad de apoyo a las políticas públicas en ciencia y tecnología.

### **4.3 Planteamiento del Modelo de Regresión Logística.**

Sea una muestra aleatoria con  $k$  variables independientes y  $n$  datos  $(y_i, x_{1i}, x_{2i}, \dots, x_{ki})$ , con  $i = 1, \dots, n$ .

Sea  $Y$  la variable aleatoria dependiente binaria, con distribución Bernoulli, con media  $\mu=p$  y varianza  $\sigma^2=pq$ , donde  $p = P(Y=1)$  es la probabilidad de éxito y  $q = 1-p$  la probabilidad de fracaso.

Sea  $OR$  el *Odds Ratio* [11], llamado también *Momios* [23] o *Cociente de probabilidades* [18], definido como la posibilidad de que un evento (éxito) se presente en un grupo de población frente al riesgo de que ocurra en otro (fracaso). Su expresión algebraica es la siguiente:

$$OR = \frac{p}{1-p}$$

Si se relacionan los valores de  $p$  con su correspondiente momio  $OR = \frac{p}{1-p}$ , se puede observar que, si bien la probabilidad  $p$  se ubica en el intervalo  $[0,1]$ , el momio toma como valor mínimo 0, pero no está acotado por el extremo superior. Este cociente se interpreta de la siguiente manera.

En un evento deportivo donde se enfrentan dos equipos se dice que el equipo A tiene 9 oportunidades contra 4 de ganarle al equipo B. Los momios a favor de A se describen como una razón: 9 a 4. En términos de probabilidades se diría que el equipo A tiene una probabilidad de  $9/(9+4)=0.6923$  de ganar el partido. Otra forma de calcular la probabilidad  $p$  de que el equipo A sea el ganador es mediante la expresión de  $OR$ , como se muestra a continuación:

$$OR = \frac{p}{1-p} = \frac{\frac{9}{13}}{\frac{4}{13}} = 2.25$$

$$p = 2.25 \cdot (1 - p)$$

$$p \cdot (1 + 2.25) = 2.25$$

$$p = \frac{2.25}{3.25} = 0.6923$$

En general:

Si  $p=0.5$ , entonces  $OR = \frac{p}{1-p} = \frac{0.5}{1-0.5} = 1$ , es decir que se tiene un éxito por cada fracaso (1 a 1).

Si  $p=0.75$ , entonces  $OR = \frac{p}{1-p} = \frac{0.75}{1-0.75} = 3$ , es decir que se tienen tres éxitos por cada fracaso (3 a 1).

Si  $p=0.25$ , entonces  $OR = \frac{p}{1-p} = \frac{0.25}{1-0.25} = \frac{1}{3}$ , es decir que se tiene un éxito por cada tres fracasos (1 a 3).

#### 4.3.1 Regresión logística.

Sea  $p=P(Y=1)$  la probabilidad de éxito. Por otro lado, la función logística  $g$  es:

$$g = \frac{1}{1 + e^{-t}}$$

Donde:

$e^t$  es la función exponencial elevada a la  $-t$ .  
 $t$  es una variable continua.

En particular, aplicando esta función a la probabilidad de éxito descrita anteriormente y expresada en función de las variables independientes, se obtiene la siguiente ecuación:

$$p_i = \frac{1}{1 + e^{-\sum_{j=0}^k \alpha_j X_j}} \quad \text{con } j = 0, 1, \dots, k$$

Donde:

$p_i$  es la probabilidad de éxito.  
 $e$  es la función exponencial.  
 $\alpha_j$  son los parámetros a estimar.  
 $k$  es el número de parámetros a estimar.

Sustituyendo la transformación logística en los momios y aplicando logaritmo natural se obtiene la siguiente expresión con mejores posibilidades de manejo algebraico:

$$\ln(OR) = \ln\left(\frac{p}{1-p}\right) = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_k X_k$$

Así, la ecuación estimada es de la forma:

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right) = \hat{\alpha}_0 + \hat{\alpha}_1 X_1 + \hat{\alpha}_2 X_2 + \dots + \hat{\alpha}_k X_k$$

Cabe mencionar que, cuando se sustituye un dato de interés ( $X_{1j}, \dots, X_{kj}$ ) en el término de la derecha, su resultado será un número real  $Y'$  con valores en el intervalo  $(-\infty, \infty)$ . Esto es,

$$-\infty < Y' = \ln\left(\frac{\hat{p}}{1 - \hat{p}}\right) < \infty$$

El modelo permite hacer una interpretación de los signos de los coeficientes. Así, el signo positivo implica que un incremento de una unidad en la variable independiente considerada implica incremento en la proporción del valor del coeficiente sobre el OR, y por lo tanto en la probabilidad de éxito del evento. Equivalentemente, el signo negativo en el parámetro implica decremento en el OR y en la probabilidad de éxito.

### 4.3.2 Estimación del modelo logístico.

Sea una muestra aleatoria con  $n$  observaciones independientes de parejas ordenadas  $(X_i, Y_i)$  donde  $Y_i$  puede tomar el valor 1 con probabilidad  $p_i$  y 0 con probabilidad  $1-p_i$ . La función de probabilidad asociada a cada  $Y_i$  se distribuye como una Bernoulli de la siguiente forma:

$$P(Y_i = 1) = p_i^{y_i} \cdot (1 - p_i)^{1-y_i}$$

Este modelo requiere estimar a los parámetros  $\alpha_i$ . Para ello, se recurre al método de máxima verosimilitud, el cual consiste en encontrar aquellos parámetros que maximizan la probabilidad de la muestra observada. En otras palabras, aquellos parámetros que hacen más creíble o verosímiles los valores observados en la muestra [23].

La función de verosimilitud es la misma función de distribución de probabilidad, pero expresada en relación con los parámetros, en vez de con las variables independientes [18]. El valor de  $\alpha_i$  que maximiza a la función de verosimilitud, también es el mismo valor que maximiza a su logaritmo [23]. De esa manera, el proceso para encontrar  $\alpha_i$  que maximiza al logaritmo de los momios es mediante el cálculo de la primera derivada, igualándola a cero para encontrar sus raíces y posteriormente verificando en la segunda derivada si las raíces encontradas representan un máximo o un mínimo.

En este caso, se cuenta con  $n$  observaciones independientes  $(y_1, y_2, \dots, y_n)$ , cuya función de probabilidad conjunta o de verosimilitud referida al vector de parámetros que se pretenden estimar es el producto de las funciones de probabilidad marginales (Bernoulli o Binomial al tener repeticiones de valores de  $Y$ ):

$$L(\alpha) = \prod_{i=1}^n p_i^{y_i} \cdot (1 - p_i)^{1-y_i}$$

Al aplicar logaritmo se transforma en:

$$\ln(L(\alpha)) = \sum_{i=1}^n [y_i \cdot \ln(p_i) + (1 - y_i) \cdot \ln(1 - p_i)]$$

Desarrollando los términos de la sumatoria, agrupándolos y sustituyendo la transformación logística en donde corresponde, la expresión anterior se puede expresar también de la siguiente forma:

$$\ln(L(\alpha)) = \sum_{i=1}^n [y_i(\alpha_0 + \alpha_1 X_{1i}) - \ln(1 + e^{\alpha_0 + \alpha_1 X_{1i}})]$$

A continuación, se deriva el logaritmo natural de la función de verosimilitud respecto de cada uno de los parámetros que intervienen en el modelo y se iguala a cero para encontrar sus raíces. Esto es:

$$\frac{\partial \ln(L(\alpha))}{\partial \alpha_i} = \sum_{i=1}^n [y_i X_i - \left(\frac{1}{1 + e^{\alpha_0 + \alpha_1 X_{1i}}}\right) X_i e^{\alpha_0 + \alpha_1 X_{1i}}]$$

$$\frac{\partial \ln(L(\alpha))}{\partial \alpha_i} = \sum_{i=1}^n [y_i X_i - p_i X_i] = 0$$

$$\frac{\partial \ln(L(\alpha))}{\partial \alpha_i} = \sum_{i=1}^n X_i [y_i - p_i] = 0$$

Este sistema de ecuaciones expresadas para los diferentes parámetros del modelo implica expresiones no lineales y, por lo tanto, no se puede resolver por mínimos cuadrados ordinarios. Para su solución se suele recurrir entre otros procedimientos, como mínimos cuadrados ponderados. Sin embargo, la mayoría de los paquetes estadísticos, como el SPSS, utilizan métodos numéricos para aproximar la solución, entre ellos el método más usual es el de Newton-Rapson [23], el cual se comenta brevemente a continuación.

Se supone que se pretende resolver una ecuación expresada por la función genérica  $f(r) = 0$ , donde  $r$  es la raíz o solución. El método parte de desarrollar la función en serie de Taylor y retener la parte lineal.

$$f(r) = f(x_{n-1}) + (r - x_{n-1})f'(x_{n-1}) + \dots$$

Se iguala a cero la expresión, pues se parte de que se cumple  $f(r) = 0$  para identificar el valor de  $r$ .

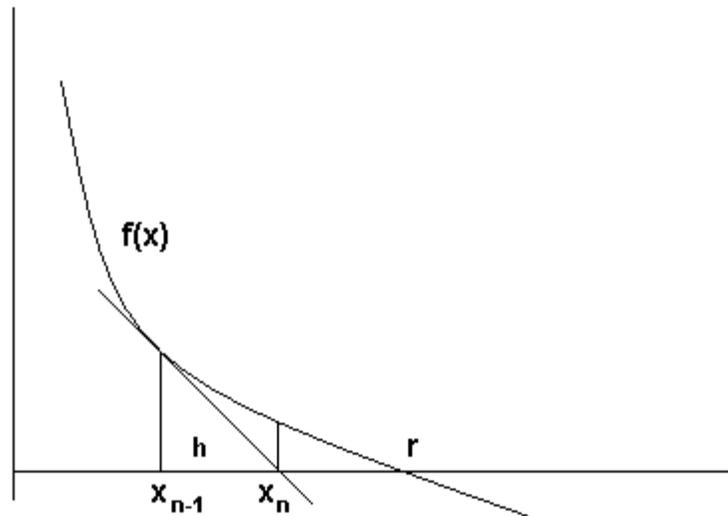
$$0 = f(x_{n-1}) + (r - x_{n-1})f'(x_{n-1}) + \dots$$

La raíz  $r$  queda aproximada por un valor inicial y se define una fórmula iterativa que converge a la raíz.

$$r \approx x_n = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}$$

La interpretación geométrica del método lleva a adoptar el valor de la raíz aproximada  $x_n$  en cada fase como el punto en donde la recta tangente definida en el punto de la aproximación anterior  $x_{n-1}$  intersecta al eje de las abscisas. El nuevo punto sirve de base para una nueva iteración.

**Figura 7. Interpretación geométrica del método de Newton-Rapson**



El proceso se interrumpe porque la diferencia en valor absoluto entre  $f(x_n)$  y el cero es menor que un valor dado, o porque se cumple cierto número de iteraciones [23].

#### 4.3.3 Ajuste del Modelo Logit [23].

Para probar que el Modelo Logit estimado que contiene a  $k$  variables independientes es estadísticamente significativo, se utiliza el estadístico *Devianza* o *Deviance*, el cual sigue una distribución  $\chi^2_{n-k-1}$ . Asimismo, se debe considerar un nivel de significancia, generalmente  $\alpha=0.05$  para el contraste de hipótesis. A continuación se detalla el método de ajuste con el estadístico Deviance.

Se parte de la comparación de los valores observados de la variable dependiente y de los valores predichos con y sin la presencia de la variable o variables explicativas. Para comprender mejor el concepto, es conveniente considerar que los valores observados son predicciones efectuadas a partir de un modelo saturado. Esto es, un modelo que contiene tantos parámetros como observaciones. Un ejemplo más

simple de un modelo saturado es definir un modelo regresión simple cuando se tienen únicamente dos puntos.

La comparación de los valores predichos y observados a través de la función de verosimilitud se basa en la expresión:

$$D = -2 \ln \left( \frac{\text{Verosimilitud del modelo ajustado}}{\text{Verosimilitud del modelo saturado}} \right)$$

El cociente dentro del paréntesis se conoce como razón de verosimilitud (likelihood ratio). El logaritmo natural de este cociente, multiplicado por (-2), se utiliza para propósitos de verificación de hipótesis. Esta prueba se conoce como la prueba de la razón de verosimilitud y la estadística D es llamada la Devianza (Deviance) y se distribuye como una  $\chi^2$  con grados de libertad dados por la diferencia entre el número de parámetros del modelo completo, menos el número de parámetros en el modelo reducido. La Devianza cumple entonces el mismo papel que la suma de cuadrados de los residuales en la regresión lineal. La Devianza se puede expresar alternativamente como la siguiente suma:

$$D = -2 \sum_{i=1}^n [y_i \ln(\hat{p}_i) + (1 - y_i) \ln(1 - \hat{p}_i)]$$

Cabe notar que como los valores de  $p_i$  son menores que 1, sus logaritmos son negativos, por lo que la Devianza siempre es positiva. Además, como los valores de Y son solamente 0 o 1, solo interviene uno de los términos de la suma.

Otra manera de medir la bondad de ajuste de la estimación del modelo consiste en elaborar la *Tabla de Clasificación*, la cual mide el número de datos clasificados correctamente entre los datos observados y los pronosticados, dada la sustitución de cada registro en la ecuación resultante.

**Tabla 6. Esquema de la Tabla de Clasificación [18].**

	<i>Pronosticado</i>		<b>Porcentaje correcto</b>
	0	1	
<i>Observado</i>			
0	$a_{00}$	$a_{01}$	$100 \cdot \frac{a_{00}}{a_{00} + a_{01}}$
1	$a_{10}$	$a_{11}$	$100 \cdot \frac{a_{11}}{a_{10} + a_{11}}$
<b>Porcentaje global</b>			$100 \cdot \frac{a_{00} + a_{11}}{\sum_{i,j=0}^1 a_{ij}}$

Dado que la sustitución de datos en el modelo estimado arrojará probabilidades de éxito  $p$  con valores en el intervalo cerrado entre 0 y 1, se debe establecer un punto de corte para clasificar las probabilidades estimadas como favorables o desfavorables. En general, el punto de corte es el punto medio del intervalo, es decir, 0.500 [18].

#### 4.3.4 Grado de contribución de las variables de forma individual [23].

Para comprobar la validez estadística de cada una de las variables consideradas en el modelo, se utiliza al estadístico de Wald, el cual se basa en el cociente entre el coeficiente de regresión y el error estándar de regresión.

La estadística de Wald se obtiene como el cociente del cuadrado del estimador entre su varianza estimada. Esta estadística se distribuye asintóticamente como Chi cuadrada con grados de libertad igual al número de parámetros estimados.

$$W(\hat{\alpha}_i) = \frac{\hat{\alpha}_i^2}{V(\hat{\alpha}_i)}$$

Y para la constante se tiene:

$$W(\hat{\alpha}_i) = \frac{\hat{\alpha}_0^2}{n\hat{p}_0(1 - \hat{p}_0)}$$

En este caso 1. Así, los resultados del cálculo del estadístico de Wald se contrastan con una  $X_{1, gdl}^2$  a un nivel de significancia dado, generalmente  $\alpha=0.05$ .

Las hipótesis nula y alternativa respectivamente son:

$H_0: \alpha_i = 0$  El parámetro  $\alpha_i$  no es significativamente estadístico

$H_1: \alpha_i \neq 0$  El parámetro  $\alpha_i$  es significativamente estadístico

La regla de decisión es: si el valor de la probabilidad asociada al estadístico de Wald es mayor que el nivel de significancia se rechaza  $H_0$ ; esto es:

Si P-Value (Wald)  $> \alpha$ , se rechaza  $H_0$ ; en otro caso, no se rechaza  $H_0$ .

#### 4.3.5 Interpretación de los coeficientes del modelo [23].

En el modelo lineal los coeficientes del modelo se interpretan como el incremento que experimenta la variable dependiente por cualquier cambio unitario de la variable independiente de referencia, cuando las demás variables en el modelo permanecen constantes. En el caso del modelo de regresión logística no resulta fácil interpretar

los coeficientes en término de las probabilidades que se modelan. Resulta más simple la interpretación si ésta se refiere a su forma linealizada por los Logits.

Partiendo del modelo en función de Logits con probabilidad de éxito  $p_0$  asociada al valor actual de la variable independiente, se tiene la siguiente ecuación:

$$\ln\left(\frac{p_0}{1-p_0}\right) = a_0 + \alpha_1 x_1$$

A continuación, se define un incremento unitario para la variable independiente y se designa por  $p_1$  su probabilidad asociada.

$$\ln\left(\frac{p_1}{1-p_1}\right) = a_0 + \alpha_1(x_1 + 1)$$

Se toma la diferencia de ambas expresiones

$$\ln\left(\frac{p_1}{1-p_1}\right) - \ln\left(\frac{p_0}{1-p_0}\right) = a_0 + \alpha_1(x_1 + 1) - (a_0 + \alpha_1 x_1)$$

En el lado izquierdo se tiene que la diferencia de logaritmos de los momios es equivalente al logaritmo del cociente de momios y en el lado derecho se concluye con el coeficiente de  $X_1$ .

$$\ln\left[\frac{\frac{p_1}{1-p_1}}{\frac{p_0}{1-p_0}}\right] = \alpha_1(x_1 + 1 - x_1) = \alpha_1$$

Así, el coeficiente de  $X_1$  se interpreta como la razón entre momios asociada a unidades consecutivas de la variable independiente.

## V RESULTADOS ESTADÍSTICOS

En este capítulo se realiza una descripción de los hallazgos obtenidos luego de realizar varios ejercicios estadísticos con datos de la ENPECYT2013. Posteriormente se explica la elección del modelo estadístico definitivo y se lleva a cabo un análisis de sus resultados. El paquete estadístico utilizado para llevar a cabo este ejercicio fue SPSS versión 20.

### 5.1 Resultados de los ejercicios estadísticos.

Para explicar a la variable dependiente *apoyo a políticas públicas en ciencia y tecnología*, en principio se eligieron 25 variables independientes, agrupadas en cuatro grandes rubros: sociodemográficas, consulta de medios, conocimientos y percepción. La lista de variables se presentó en el capítulo III, punto 3.5.1 *Variables del Modelo*.

El primer ejercicio estadístico reportó que el modelo con las 25 variables seleccionadas contenía 80.3% de información correcta, de acuerdo con los resultados de la tabla de clasificación resultante. De lo anterior, se concluye que, de manera general, es aceptable el modelo con esta selección, ya que clasifica correctamente a la mayor parte de los datos, como se puede apreciar en la siguiente tabla (Ver Anexo 1).

**Tabla 7. Tabla de clasificación considerando 25 variables dependientes.**

Observados	Predichos		
	No apoya	Si apoya	Porcentaje
No apoya	126	510	19.8
Si apoya	54	2,176	97.6
			80.3

De acuerdo con los resultados del estadístico de *Wald*, 13 variables de percepción, así como las variables *edad*, *consulta de CyT en periódicos* y *consulta de CyT en radio* no son estadísticamente significativas en este modelo, por lo que fueron apartadas del análisis para realizar otros ejercicios estadísticos con las variables restantes. A continuación, se presentan los resultados que explican lo anterior.

**Tabla 8. Resultados estadísticos del modelo Logit con 25 variables dependientes.**

	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
							Lower	Upper
Sexo	.296	.100	8.788	1	.003	1.345	1.106	1.635
Edad	-.001	.003	.179	1	.672	.999	.992	1.005
Escolaridad	.038	.013	9.168	1	.002	1.039	1.013	1.065
CYT en TV	.008	.004	4.960	1	.026	1.008	1.001	1.015
CYT en radio	.001	.006	.400	1	.841	1.001	.990	1.013
CYT en periódico	.005	.004	1.116	1	.291	1.005	.996	1.013
CYT en revistas	.007	.003	5.367	1	.021	1.007	1.001	1.014
Calificación	.012	.004	11.327	1	.001	1.012	1.005	1.019
Gracias a la ciencia y la tecnología habrá más oportunidades para las próximas generaciones	.069	.069	1.003	1	.317	1.072	.936	1.227
Los descubrimientos tecnológicos tarde o temprano destruirán el planeta	.008	.055	.021	1	.885	1.008	.906	1.122
Debido a sus conocimientos, los investigadores científicos tienen un poder que los hacen peligrosos	-.074	.059	1.555	1	.212	.929	.827	1.043
Los científicos deben ser libres de llevar a cabo sus investigaciones a su antojo, siempre y cuando las hagan bajo reglas	.155	.060	6.734	1	.009	1.168	1.039	1.313
En México debería haber más gente trabajando en investigación y desarrollo tecnológico	.142	.076	3.452	1	.063	1.152	.992	1.338
Los científicos y los empresarios deberían cooperar más entre sí	-.069	.086	.654	1	.419	.933	.789	1.104
Los científicos deberían interesarse más en patentar sus investigaciones y en el uso que se les dé	.084	.065	1.662	1	.197	1.087	.957	1.234
Debería haber más mujeres dedicadas a la investigación científica en nuestro país	.171	.073	5.420	1	.020	1.186	1.027	1.370
El gobierno debe impulsar que las personas participen en debates sobre la asignación de presupuesto para ciencia y tecnología	.002	.071	.000	1	.982	1.002	.872	1.151
La ciencia tiene una imagen muy negativa para la sociedad	.141	.058	5.980	1	.014	1.151	1.028	1.289
Confiamos demasiado en la fe y muy poco en la ciencia	.041	.066	.394	1	.530	1.042	.916	1.187
2 La aplicación de la ciencia hace que nuestro modo de vida cambie demasiado rápido	.009	.071	.018	1	.894	1.010	.878	1.161
3 Algunos números son de la suerte	-.051	.065	.609	1	.435	.950	.837	1.080
4 El desarrollo tecnológico origina una manera de vivir artificial y deshumanizada	.193	.057	11.265	1	.001	1.213	1.083	1.357
5 Algunos de los objetos voladores no identificados que se han reportado, son en realidad vehículos espaciales de otras	.017	.054	.095	1	.758	1.017	.915	1.130
6 Algunas personas poseen poderes psíquicos	-.017	.061	.080	1	.777	.983	.871	1.109
7 Existen medios adecuados para el tratamiento de enfermedades que la ciencia no reconoce (acupuntura, quiropráctica,	.093	.063	2.159	1	.142	1.097	.969	1.242
Constant	-2.567	.378	46.241	1	.000	.077		

Como se puede apreciar en la tabla anterior, las variables resaltadas que no son estadísticamente significativas reportaron valores de Wald muy bajos, asociados a su correspondiente probabilidad muy por encima del nivel de significancia establecido de 0.05. Por ejemplo, la variable *edad* no es significativa estadísticamente dentro de este modelo, pues su valor calculado fue sig.= 0.672 mayor que 0.05. Por lo tanto, esta variable no explica al modelo.

## 5.2 Modelo resultante.

Luego de varios ejercicios de exploración se encontró el que mejor ajusta al modelo, basado en el uso de nueve variables independientes. De sus resultados se reporta que 80.2% de los casos fueron correctamente clasificados y todas las variables independientes significativas estadísticamente ( $\alpha = 0.05$ ) (Ver Anexo 2).

**Tabla 9** Tabla de clasificación del modelo resultante<sup>a</sup>

<i>Observado</i>	<i>Pronosticado</i>		
	No apoya	Si apoya	Porcentaje correcto
No apoya	119	517	18.7
Si apoya	49	2,172	97.8
<b>Porcentaje global</b>			<b>80.2</b>

a. El valor de corte es .500

Así, las variables que componen el modelo son:

Variable dependiente:

- Apoyo a políticas públicas en ciencia y tecnología

Variables independientes:

- Sexo
- Escolaridad
- Consulta de CyT en televisión
- Consulta de CyT en revistas
- Calificación
- Los científicos deben ser libres de llevar a cabo sus investigaciones a su antojo, siempre y cuando las hagan bajo reglas éticas.
- Debería haber más mujeres dedicadas a la investigación científica en nuestro país.
- La ciencia tiene una imagen muy negativa para la sociedad.

- El desarrollo tecnológico origina una manera de vivir artificial y deshumanizada.

**Tabla 10. Resultados estadísticos del modelo resultante.**

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 <sup>a</sup>						
Sexo	.316	.099	10.199	1	.001	1.371
Escolaridad	.046	.012	15.051	1	.000	1.047
Consulta de CyT en televisión	.009	.004	6.328	1	.012	1.009
Consulta de CyT en revistas	.008	.003	6.508	1	.011	1.008
Calificación	.015	.003	18.165	1	.000	1.015
Los científicos deben ser libres de llevar a cabo sus investigaciones a su antojo, siempre y cuando las hagan bajo reglas éticas.	.179	.056	10.225	1	.001	1.196
Debería haber más mujeres dedicadas a la investigación científica en nuestro país.	.241	.062	14.877	1	.000	1.272
La ciencia tiene una imagen muy negativa para la sociedad.	.172	.055	9.933	1	.002	1.188
El desarrollo tecnológico origina una manera de vivir artificial y deshumanizada.	.216	.052	17.228	1	.000	1.241
Constant	-2.475	.254	94.774	1	.000	.084

En la tabla 10 se presentan las estimaciones de los parámetros  $\alpha_i$  en correspondientes a cada variable independiente, así como el estadístico Wald y su probabilidad asociada (Sig.), cuyos valores son todos menores que el nivel de significancia predefinido (0.05). Asimismo, se ofrece el valor de los momios en la columna Exp(a).

De lo anterior, la ecuación resultante es:

$$Y' = \ln\left(\frac{\hat{p}}{1-\hat{p}}\right) = -2.476 + 0.316 \cdot \text{Sexo} + 0.241 \cdot (\text{más mujeres en CyT}) + 0.216 \cdot (\text{vida artificial}) + 0.179 \cdot (\text{reglas éticas}) + 0.172 \cdot (\text{imagen negativa}) + 0.046 \cdot \text{Escolaridad} + 0.015 \cdot \text{Calificación} + 0.009 \cdot \text{CyT TV} + 0.008 \cdot \text{CyT Revistas}$$

Donde  $Y'$  es el logaritmo natural del *Odds Ratio* estimado de la variable *apoyo a políticas públicas en ciencia y tecnología*.

De los resultados obtenidos por este ejercicio estadístico, se realizó el cálculo de la elación de momios que es de utilidad para realizar el análisis de los resultados. Para ello, se realizó la estimación de la probabilidad asociada a cada variable a partir de la columna Exp(a), que es el resultado numérico de los momios; esto es:

$$\frac{p}{1-p} = e^{\alpha} = \text{Exp}(\alpha)$$

Despejando p:

$$p = \frac{\text{Exp}(\alpha)}{1 + \text{Exp}(\alpha)}$$

**Tabla 11. Probabilidades individuales y momios asociados a cada variable.**

Variable	Exp(a)	p	1-p	Diferencia %
Sexo	1.371	0.578	0.422	37.1
Escolaridad	1.047	0.512	0.488	4.7
Consulta de CyT en televisión	1.009	0.502	0.498	0.9
Consulta de CyT en revistas	1.008	0.502	0.498	0.8
Calificación	1.015	0.504	0.496	1.5
Los científicos deben ser libres de llevar a cabo sus investigaciones a su antojo, siempre y cuando las hagan bajo reglas éticas.	1.196	0.545	0.455	19.6
Debería haber más mujeres dedicadas a la investigación científica en nuestro país.	1.272	0.560	0.440	27.2
La ciencia tiene una imagen muy negativa para la sociedad.	1.188	0.543	0.457	18.8
El desarrollo tecnológico origina una manera de vivir artificial y deshumanizada.	1.241	0.554	0.446	24.1

### 5.3 Análisis de los resultados.

Es importante notar que, salvo la constante cuyo signo es negativo, todas las demás estimaciones son de signo positivo, lo cual indica que la probabilidad resultante de apoyar a las políticas públicas en ciencia y tecnología de cada variable tendrá una relación directa con la escala de medición de cada una de ellas.

Tomando en cuenta el peso que tienen los parámetros estimados  $\hat{\alpha}_i$  en el modelo, la variable más influyente es Sexo ( $\hat{\alpha} = 0.316$ ) la cual, de acuerdo con su codificación<sup>5</sup>, indica que los hombres son más propensos a apoyar a las políticas públicas en CyT que las mujeres, como lo indica el signo positivo del parámetro estimado. Por otro lado, al interpretar el valor de los momios, que es 1.371, se aprecia que la probabilidad de apoyo a políticas públicas es 37.1% mayor que la probabilidad de no apoyarlas.

Después de la variable sexo, las variables independientes con mayor injerencia son las que están relacionadas con la percepción pública.

Así, de las variables de percepción, la que se ubica en primer lugar es *debería haber más mujeres dedicadas a la investigación científica en nuestro país* ( $\hat{\alpha} = 0.241$ ).

<sup>5</sup> Codificación de la variable Sexo: 0 = Mujeres, 1 = Hombres.

Esta variable se interpreta como un deseo favorable de la sociedad por una mayor participación de las mujeres en actividades científicas y tecnológicas, reconociendo que dicha contribución es baja o insuficiente, y su correspondiente *Odds Ratio* es de 1.272, lo cual significa que la probabilidad de que las personas que opinan favorablemente de esta variable son 27.2% más propensas a apoyar a las políticas públicas que de no hacerlo.

Por su parte, la variable *el desarrollo tecnológico origina una manera de vivir artificial y deshumanizada* ( $\hat{\alpha} = 0.216$ ) tiene un peso relevante sobre políticas públicas, lo cual es aparentemente contradictorio, debido a que expresa reservas ante el avance tecnológico, incluso una contraposición al mismo. Sin embargo, la gente confía en el apoyo a las políticas públicas en ciencia y tecnología con una probabilidad 24.1% mayor que la probabilidad de no apoyarlas.

La variable *los científicos deben ser libres de llevar a cabo sus investigaciones a su antojo, siempre y cuando las hagan bajo reglas éticas* ( $\hat{\alpha} = 0.179$ ) tiene un poco menos de impacto sobre políticas públicas de ciencia y tecnología que las anteriores variables de percepción pública, siendo su *Odds Ratio* de 1.195, es decir, la probabilidad de apoyar políticas públicas en ciencia y tecnología mediante esta variable es 19.5% mayor que de no apoyarlas. Cabe mencionar que esta variable representa una opinión favorable de la sociedad hacia el desarrollo de la ciencia, siempre que se observen reglas éticas en el mismo.

Finalmente, la variable de percepción *la ciencia tiene una imagen muy negativa para la sociedad* ( $\hat{\alpha} = 0.172$ ) puede interpretarse en dos sentidos. Uno es que cada individuo de la sociedad considere a la ciencia con mala imagen para todos, incluido él mismo, pero también se puede considerar que la opinión de esa persona está en relación con lo que considera del resto de la sociedad, sin incluirse. Su *Odds Ratio* es de 1.188. Así, la probabilidad de que las políticas públicas en ciencia y tecnología sean favorecidas mediante esta variable es 18.8% mayor que la probabilidad de no apoyarlas.

Las variables *calificación* (conocimientos científicos) y *escolaridad* (nivel educativo), aunque sean estadísticamente significativas, tienen poco peso en el apoyo a las políticas públicas en ciencia y tecnología. Es de observar que ambas son de signo positivo, lo cual indica que conforme sea mayor la escolaridad y los conocimientos, mayor será el apoyo a tales políticas públicas.

Así, la *escolaridad* ( $\hat{\alpha} = 0.046$ ), medida en años de escuela concluidos, teóricamente debería ser un factor decisivo y favorable para el apoyo a las políticas públicas en este tema. Sin embargo, de acuerdo con su *Odds Ratio* de 1.047, la probabilidad de que las personas conforme es mayor su nivel de estudios apoyen a las políticas públicas en ciencia y tecnología es solamente 4.7% mayor que de no apoyarlas.

De la misma manera que con *escolaridad*, la variable *calificación* ( $\hat{\alpha} = 0.015$ ) reporta un resultado débil comparado con el esperado, ya que su *Odds Ratio* de 1.015 indica que la probabilidad de que conforme sea mayor el conocimiento de ciencia y

tecnología, se espera que el apoyo a políticas públicas en el tema sea solamente 1.5% mayor que el de no apoyarlas.

Finalmente, las variables de consulta en temas de ciencia y tecnología en los medios de comunicación masiva se pueden considerar neutrales en cuanto a su capacidad de explicación de la variable dependiente, pese a que sean estadísticamente significativas.

De esta manera, la variable *consulta de ciencia y tecnología en televisión* ( $\hat{\alpha} = 0.009$ ), reporta un *Odds Ratio* de 1.009, es decir, la probabilidad de que la gente con mayor nivel de consulta de estos temas en televisión apoye a políticas públicas en ciencia y tecnología es solamente 0.9% mayor que de no apoyarlas. De forma equivalente, los resultados la *consulta de ciencia y tecnología en radio* ( $\hat{\alpha} = 0.008$ ) indican que la probabilidad de apoyo es mayor solamente en 0.8% que de no apoyo a las políticas públicas en ciencia y tecnología.

De acuerdo con los resultados encontrados, se concluye que, aunque las variables utilizadas para el modelo son estadísticamente significativas y el modelo en general es aceptable (80.2 de datos clasificados correctamente), su injerencia para describir la probabilidad de apoyar a las políticas públicas en ciencia y tecnología es muy baja. El mayor incremento por cambio en valor de variable se observa en *sexo*, con 37.1%, mientras que en variables de mucha relevancia como son *conocimientos* y *escolaridad* presentan momios con probabilidad prácticamente de 0.5, lo que significa que la varianza es máxima y, por lo tanto, el nivel de conocimientos y de escolaridad resultan irrelevantes.

Esto lleva a plantear la posibilidad de analizar el desempeño de los medios de comunicación masiva y del sistema escolarizado, pues de ambos surge la información sobre ciencia y tecnología que, en el caso de este ejercicio, no parece que influyan de forma decidida en la percepción pública de la ciencia y la tecnología.

## **VI CONCLUSIONES Y RECOMENDACIONES**

Este capítulo recoge los resultados obtenidos en el ejercicio estadístico anteriormente analizado y ofrece una breve discusión sobre ellos. Asimismo, da cuenta tanto del cumplimiento de los objetivos, como del incumplimiento de las hipótesis planteadas al inicio de este trabajo, mediante el reporte de comparaciones internacionales del nivel y la calidad educativa de México y el análisis de algunos resultados de la ENPECYT 2013. Finalmente se hacen recomendaciones en materia de políticas públicas orientadas al mejoramiento de los componentes de ciencia y tecnología en el sector educativo y en los medios de comunicación.

### **6.1 Conclusiones**

De acuerdo con los resultados estadísticos del Modelo Logit reportados en el capítulo anterior, todas las variables independientes aportan poco peso a la variable dependiente. Cabe recordar que la variable sexo es la de mayor peso dentro del modelo, seguida de las variables de percepción pública de la ciencia y la tecnología y, en último lugar, las variables de nivel educativo, conocimientos científicos y tecnológicos y consulta de estos temas en los medios.

Aun considerando un impacto bajo de las variables sexo y especialmente de percepción en el modelo, se observa que las personas que manifiestan de forma indistinta opiniones favorables o de reserva ante el avance científico y tecnológico, están dispuestas a brindar su apoyo a las políticas públicas en ciencia y tecnología.

En contraste, la formación académica, así como la consulta regular de medios de información son variables prácticamente irrelevantes para definir el perfil de las personas que apoyan a las políticas públicas en ciencia y tecnología, por lo que no se cumplen las hipótesis planteadas en el presente trabajo.

Estos resultados poco cercanos a los conceptos y modelos mencionados en este trabajo se desglosan a continuación.

### **6.2 Cumplimiento de los objetivos.**

A partir de los resultados del Modelo Logit se obtuvo la ecuación correspondiente, con lo que se identificó el conjunto de factores que inciden en el apoyo de los ciudadanos a las políticas públicas en ciencia y tecnología, cumpliendo así con el primer objetivo planteado en el Capítulo I.

El segundo objetivo, que consiste en identificar variables o factores específicos y su impacto individual en el tema de apoyo a políticas públicas, también se logró mediante los resultados estadísticos descritos en el capítulo anterior. En este sentido es importante recordar que las principales variables independientes que afectan de manera favorable a la decisión de apoyar a las políticas públicas en ciencia y tecnología son sexo y las relacionadas con la percepción pública.

La ecuación resultante de la corrida del modelo Logit permite calcular la probabilidad de apoyo a políticas públicas en ciencia y tecnología, dado un caso particular de ciudadano con perfil específico, con lo que se cumple el tercer objetivo.

Todo lo anterior permite emitir recomendaciones para fortalecer o definir políticas públicas en favor de la ciencia y la tecnología, lo cual dará cuenta del cumplimiento del objetivo 4 en las recomendaciones de política de este mismo capítulo.

### 6.3 Cumplimiento de las hipótesis.

No se cumplieron las dos hipótesis planteadas, como a continuación se explica.

**Hipótesis 1:** La formación educativa es un factor relevante en la percepción pública y en el apoyo a las políticas públicas en ciencia y tecnología.

Los resultados estadísticos antes mencionados indican que el nivel educativo es estadísticamente significativo, pero con impacto casi nulo para el apoyo a las políticas públicas en ciencia y tecnología.

Adicionalmente, cabe mencionar que las variables *escolaridad* y *conocimientos científicos básicos* están relacionadas positivamente entre sí. Es decir, a mayor escolaridad, mayor nivel de conocimientos, pero su correlación es débil ( $R^2 = 0.295$ ). De esta manera, la escolaridad no solamente es poco relevante para un mayor conocimiento en ciencia y tecnología, sino que tampoco impacta decididamente en el apoyo a las políticas públicas en dicho tema.

Lo anterior justifica la existencia de dos características básicas de la educación que inciden en la formación de las personas. Uno es el nivel educativo y otro es la calidad de la educación. En relación con el nivel educativo, de acuerdo con datos del INEGI [10], la escolaridad promedio de los mexicanos en 2010 fue de 8.6 años, equivalente a poco más de 2º de secundaria, y en 2015 se incrementó a 9.2 años, equivalente a poco más de secundaria concluida. Evidentemente no es un nivel muy alto, sobre todo al compararlo con el de otros países que son referencia para México en el tema de la ciencia y la tecnología, los cuales superan a nuestro país en más de 3 años efectivos. Por ejemplo, las naciones con mayores niveles de escolaridad, de acuerdo con la página de información mundial DBCity [6], son Noruega (12.6), Nueva Zelanda (12.5), Estados Unidos (12.4), República Checa (12.3) y Alemania (12.2), todos equivalentes a bachillerato concluido.

Por su parte, es conocido que la calidad de la educación en México está rezagada en comparación con la de muchos países. Por ejemplo, la formación de nivel básico en México claramente no es competitiva a nivel internacional, como lo demuestran los indicadores del informe PISA [16], en los que los estudiantes mexicanos de niveles primaria y secundaria se ubican entre los últimos lugares de los países en los que se aplican tres pruebas, una de ciencias, otra de comprensión de lectura y

una de matemáticas. Los países con mejor desempeño en estas evaluaciones son principalmente de Asia oriental y del norte de Europa, así como Canadá. Incluso varios países latinoamericanos obtienen mejores notas que México en este reporte.

Así, en las pruebas de ciencias, México se ubica en el lugar 58 de 70 países, siendo superado por varios latinoamericanos como Colombia (57), Costa Rica (55), Trinidad y Tobago (53), Uruguay (47), Chile (44) y Argentina (38) [16].

En comprensión de lectura, México ocupa el lugar 55 de 70, mientras que Colombia (54), Costa Rica (52), Trinidad y Tobago (51), Uruguay (46), Chile (42) y Argentina (38) obtuvieron mejores resultados [16].

Y en matemáticas, la posición de México está en el lugar 56 de 70, siendo mejores Trinidad y Tobago (53), Uruguay (51), Chile (48) y Argentina (42) [16].

Asimismo, a nivel superior solamente la Universidad Nacional Autónoma de México (UNAM), en lugar 122, y el Instituto Tecnológico de Estudios Superiores de Monterrey (ITESM), en lugar 199, están clasificadas entre las primeras 200 mejores universidades del mundo, de acuerdo con el Ranking Mundial de Universidades<sup>6</sup> (QS World University Rankings) [20], siendo la UNAM superada en Latinoamérica por la Universidad de Buenos Aires (posición 75) y la Universidad de Sao Paulo (posición 121), mientras que al ITESM lo superan la Pontificia Universidad Católica de Chile (posición 137) y la Universidad Estatal de Campiñas (posición 182). Otras universidades mexicanas que aparecen en este ranking ya no cuentan con una posición específica y son ubicadas en rangos. Así, en el rango de las posiciones entre 601 y 650 se encuentran la Universidad Anáhuac, la Universidad Iberoamericana, el Instituto Politécnico Nacional (IPN) y el Instituto Tecnológico Autónomo de México (ITAM); la Universidad Autónoma Metropolitana (UAM) se posiciona en el rango de 751 a 800 y en el rango de 801 a 850 están la Benemérita Universidad Autónoma de Puebla (BUAP), la Universidad Autónoma de Nuevo León (UANL), la Universidad Autónoma del Estado de México (UAE México), la Universidad de Guadalajara (U de G), la Universidad de las Américas de Puebla (UDLAP) y la Universidad de Monterrey [20].

Tal clasificación indica claramente que el sistema de educación superior de México, el cual incluye a licenciaturas y posgrados, en general no es verdaderamente competitivo a nivel internacional, salvo los casos de la UNAM y el ITESM.

**Hipótesis 2:** La consulta regular de información científica y tecnológica de las personas es un factor que impacta en el apoyo a las políticas públicas en ciencia y tecnología.

---

<sup>6</sup> QS World University Rankings 2018. Evalúa el desempeño académico homologado de las universidades a través de seis métricas: Reputación académica, Reputación de los empleadores, número de profesores entre estudiantes, número de citas recibidas por todos los artículos publicados por la universidad, porcentaje de profesores extranjeros y porcentaje de estudiantes extranjeros.

Como se comentó en el Capítulo II se eligieron 25 variables independientes para definir el modelo de esta tesis, entre las cuales se consideraron la consulta de información científica y tecnológica en periódicos, revistas, radio y televisión, y como resultado del primer ejercicio estadístico, solamente *consulta de CyT en televisión* y *consulta de CyT en revistas* fueron estadísticamente significativas, las demás no.

Sin embargo, estas variables significativas tienen un peso prácticamente nulo en el apoyo a las políticas públicas en ciencia y tecnología, por lo que no se pueden considerar como explicativas.

En nuestro país y en el resto del mundo, la divulgación de los alcances, desarrollo e importancia de la ciencia y la tecnología es deficiente. Los medios de comunicación masiva centran sus noticias en temas de interés general (espectáculos, deportes, internacionales, entre otros), lo cual es entendible, pues deben vender sus productos. Sin embargo, la ciencia y tecnología no forman parte de ese interés general, dado que la gente no demanda información al respecto.

**Tabla 12. Percepción del grado de contenido científico de diferentes disciplinas en México, 2013 [5]**  
Porcentaje

Disciplina	Muy científica/ Científica	Algo científica/ Nada científica	No sabe o no la conoce
Medicina	91.4	6.1	2.5
Física	81.5	13.2	5.3
Biología	79.5	14.4	6.0
Matemáticas	79.0	17.3	3.8
Astronomía	74.0	19.7	6.3
Psicología	65.4	29.9	4.8
Historia	48.9	45.9	5.2
Homeopatía	44.6	45.0	10.3
Astrología	43.2	50.3	6.5
Parasicología	40.2	49.8	10.0
Teología	37.7	51.7	10.6
Economía	37.1	56.2	6.7

No solamente hay carencia de información en ciencia y tecnología, también hay fallas en la calidad de la misma. Por ejemplo, los medios difunden programas y contenidos de pseudociencias que confunden a la gente. De acuerdo con datos de la ENPECYT 2013, de las personas que respondieron a la encuesta, 43.2 por ciento consideraron que la Astrología es muy científica o científica, mientras que 40.2 por ciento afirmaron lo mismo de la Parasicología y 37.7% de la Teología, relegando al último lugar a la Economía, con 37.1 por ciento, lo cual es lamentable dado que ésta

última disciplina es una ciencia social y las mencionadas no son científicas en lo absoluto.

Se puede concluir que la escasa y deficiente comunicación sobre temas de ciencia y tecnología en los medios redundan en mal entendimiento y confusión del público. Son pocos los medios que cuentan con secciones de divulgación de la ciencia y la tecnología, o bien que estén dedicados a la misma.

Por lo anterior, la mayoría de la gente tiene conocimiento y entendimiento ambiguos sobre los descubrimientos científicos y desarrollos tecnológicos y, aunque es usuaria de muchos desarrollos tecnológicos, como los medicamentos de nueva generación, transporte o la comunicación y demás aplicaciones que ofrecen los dispositivos electrónicos, no tiene conciencia plena de los esfuerzos que hay detrás de tales progresos.

#### **6.4 Recomendaciones de políticas públicas.**

Debido a que en México la formación académica y la consulta regular de información científica y tecnológica en los medios no son factores relevantes para el apoyo a políticas públicas en el tema y ello es contradictorio a los planteamientos teóricos nacionales e internacionales, se trazan recomendaciones agrupadas en dos temas principales: políticas en educación y políticas de divulgación.

##### **Recomendaciones para las políticas en educación.**

En primer lugar y de forma decidida, el sistema educativo nacional debe plantear como uno de sus objetivos principales impulsar el conocimiento científico y tecnológico en todos los niveles educativos, mediante dos estrategias básicas consistentes en incrementar la presencia de contenidos y materias de ciencia y tecnología, y mejorar la calidad del contenido de las ya existentes y las de nueva incorporación.

Al hacer una revisión exhaustiva de las materias de ciencia y tecnología, se deben tomar en cuenta varias consideraciones. En primer lugar, los contenidos de las materias deben ser actualizados y expresados de forma que sea comprensible para los alumnos. Muchas veces el conocimiento científico ya tiene nuevas definiciones o clasificaciones y el sistema educativo no las conoce, de manera que la enseñanza queda rezagada a los términos anteriores, sobre todo en los niveles educativos básicos.

Asimismo, cuando las ciencias en los niveles básicos se expresan con muchos tecnicismos, los alumnos difícilmente las comprenden y ello provoca que no sean de su agrado. Ello refuerza la idea de plantear nuevas formas didácticas que hagan de las ciencias materias atractivas para su estudio. De hecho, para cada materia se deben considerar la incorporación de casos prácticos fácilmente asimilables y que sean motivadores, así como anécdotas e historia de cada una de las ciencias, todo

de forma sencilla. Lo anterior, con la finalidad de hacer más tangible a la ciencia y tecnología y que provoque en los estudiantes el interés que bajo el sistema actual no se despierta.

### **Recomendaciones para las políticas de divulgación.**

La divulgación de la ciencia y la tecnología corre a cargo de diferentes actores, entre los que se encuentran los medios de comunicación masiva, tanto públicos como privados, así como las universidades y centros de investigación, el gobierno y algunas organizaciones no gubernamentales (ONG).

Como se mencionó anteriormente, los medios de comunicación masiva carecen de alcance y calidad en la divulgación de este tema. Para revertir este problema, una política pública sugerida consiste en dar incentivos fiscales a los medios que incluyan programas y contenidos de ciencia y tecnología de forma más amplia y con la calidad suficiente, avalada por expertos en el tema. También pueden otorgarse este tipo de incentivos a empresas patrocinadoras de los medios, a las cuales también se les puede dar información sobre apoyos que otorga el Gobierno Federal a través del Conacyt, la Secretaría de Economía y otras dependencias, para que las motive a desarrollar nuevos productos basados en desarrollos tecnológicos propios.

Por su parte, el Conacyt cuenta con instrumentos y programas de divulgación de la ciencia, como son las revistas *Ciencia y Desarrollo* y *Helix* (orientada a público infantil), la *Semana Nacional de Ciencia y Tecnología* y otros eventos realizados de manera esporádica. Asimismo, mediante informes específicos, da cuenta de los resultados y avances del sistema de ciencia y tecnología del país, como son el *Informe General del Estado de la Ciencia, la Tecnología y la Innovación en México* (publicación anual), informes de resultados e impacto de los fondos mixtos y sectoriales (publicados cada cierto tiempo, pero incluyendo información de un periodo amplio), casos de éxito de programas de empresas innovadoras (publicados eventualmente), entre otros.

Estos instrumentos de divulgación no cuentan con cobertura amplia, sobre todo por falta de promoción y, por lo tanto, no impactan a la sociedad en general, solo a unos pocos grupos relacionados con las universidades y centros de investigación, así como ciertas áreas gubernamentales. Se puede afirmar que son programas de autoconsumo. Ello implica realizar un diagnóstico de tales instrumentos y su correspondiente reforma, la cual deberá ir acompañada de un mayor impulso presupuestal para eliminar o al menos disminuir la exclusión social en el tema.

# ANEXO 1. EJERCICIO ESTADÍSTICO CON 25 VARIABLES

```
LOGISTIC REGRESSION VARIABLES APOYOCYT
/METHOD=ENTER SEXO EDAD ESCOLARIDAD CYTenTV CYTRADIO CYTPERIODICO CYTREVISTAS CALIFICACION S4P23_2 S4P23_6 S4P25_3 S
/PRINT=GOODFIT CORR ITER(1) CI(95)
/CRITERIA=PIN(0.05) POUT(0.10) ITERATE(20) CUT(0.5) .
```

## Logistic Regression

Notes		03-JUL-2018 14:46:07
Output Created		
Comments		
Input	Data	F:\New Folder\BASE.sav
	Active Dataset	DataSet2
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data	2858
	File	
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing
Syntax		LOGISTIC REGRESSION VARIABLES APOYOCYT /METHOD=ENTER SEXO EDAD ESCOLARIDAD CYTenTV CYTRADIO CYTPERIODICO CYTREVISTAS CALIFICACION S4P23_2 S4P23_6 S4P25_3 S4P25_7 S4P26_3 S4P26_6 S4P26_8 S4P26_10 S4P26_11 S4P28_4 S4P33_2_1 S4P33_2_2 S4P33_2_3 S4P33_2_4 S4P33_2_5 S4P33_2_6 S4P33_2_7 /PRINT=GOODFIT CORR ITER(1) CI(95) /CRITERIA=PIN(0.05) POUT(0.10) ITERATE(20) CUT(0.5).
Resources	Processor Time	00:00:00.11
	Elapsed Time	00:00:00.19

[DataSet2] F:\New Folder\BASE.sav

### Case Processing Summary

Unweighted Cases <sup>a</sup>		N	Percent
Selected Cases	Included in Analysis	2857	100.0
	Missing Cases	1	.0
	Total	2858	100.0
Unselected Cases		0	0.0
Total		2858	100.0

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
No apoya	0
Apoya	1

Block 0: Beginning Block

Iteration History<sup>abc</sup>

Iteration		-2 Log likelihood	Coefficients	
			Constant	
Step 0	1	3039.602	1.110	
	2	3029.535	1.245	
	3	3029.523	1.251	
	4	3029.523	1.251	

a. Constant is included in the model.

b. Initial -2 Log Likelihood: 3029.523

c. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.

Classification Table<sup>ab</sup>

Observed			Predicted		
			Apoyo a políticas públicas en Ciencia y Tecnología		Percentage Correct
			No apoya	Apoya	
Step 0	Apoyo a políticas públicas en Ciencia y Tecnología	No apoya Apoya	0 0	636 2221	0.0 100.0
Overall Percentage					77.7

a. Constant is included in the model.

b. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	1.251	.045	773.166	1	.000	3.492

Variables not in the Equation

			Score	df	Sig.
Step 0	Variables	SEXO	17.644	1	.000
		EDAD	28.659	1	.000
		ESCOLARIDAD	151.580	1	.000
		CYTenTV	55.196	1	.000
		CYTRADIO	7.040	1	.008
		CYTPERIODICO	31.748	1	.000
		CYTREVISTAS	43.234	1	.000
		CALIFICACION	190.679	1	.000
		S4P23_2	111.648	1	.000
		S4P23_6	40.948	1	.000
		S4P25_3	27.565	1	.000
		S4P25_7	104.160	1	.000
		S4P26_3	169.852	1	.000
		S4P26_6	126.154	1	.000
		S4P26_8	137.459	1	.000
		S4P26_10	116.787	1	.000
		S4P26_11	145.490	1	.000
		S4P28_4	87.710	1	.000
		S4P33_2_1	20.565	1	.000
		S4P33_2_2	100.510	1	.000
		S4P33_2_3	2.006	1	.157
		S4P33_2_4	104.884	1	.000
		S4P33_2_5	35.463	1	.000
		S4P33_2_6	12.364	1	.000
		S4P33_2_7	59.469	1	.000
Overall Statistics			381.860	25	.000

Block 1: Method = Enter

Iteration History<sup>abc,d</sup>

Iteration		-2 Log likelihood	Coefficients						
			Constant	SEXO	EDAD	ESCOLARIDAD	CYTenTV	CYTRADIO	CYPERIODICO
Step 1	1	2724.929	-1.777	.164	-.001	.027	.003	.001	.002
	2	2671.550	-2.441	.270	-.001	.037	.006	.001	.004
	3	2669.776	-2.564	.295	-.001	.038	.008	.001	.004
	4	2669.770	-2.567	.296	-.001	.038	.008	.001	.005
	5	2669.770	-2.567	.296	-.001	.038	.008	.001	.005

CYTREVISTAS	CALIFICACION	S4P23_2	S4P23_6	S4P25_3	S4P25_7	S4P26_3
.003	.008	.056	.017	-.032	.100	.111
.006	.011	.068	.010	-.065	.145	.138
.007	.012	.069	.008	-.074	.155	.142
.007	.012	.069	.008	-.074	.155	.142
.007	.012	.069	.008	-.074	.155	.142

S4P26_6	S4P26_8	S4P26_10	S4P26_11	S4P28_4	S4P33_2_1	S4P33_2_2
-.049	.078	.107	.019	-.106	.019	.012
-.067	.085	.159	.006	.136	.037	.009
-.069	.084	.171	.002	.141	.041	.009
-.069	.084	.171	.002	.141	.041	.009
-.069	.084	.171	.002	.141	.041	.009

S4P33_2_3	S4P33_2_4	S4P33_2_5	S4P33_2_6	S4P33_2_7
-.028	.137	.011	-.008	.065
-.047	.186	.015	-.014	.088
-.051	.193	.017	-.017	.093
-.051	.193	.017	-.017	.093
-.051	.193	.017	-.017	.093

- a. Method: Enter
- b. Constant is included in the model.
- c. Initial -2 Log Likelihood: 3029.523
- d. Estimation terminated at iteration number 5 because parameter estimates changed by less than .001.

Omnibus Tests of Model Coefficients

Step		Chi-square	df	Sig.
Step 1	Step	359.752	25	.000
	Block	359.752	25	.000
	Model	359.752	25	.000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	2669.770 <sup>a</sup>	.118	.181

- a. Estimation terminated at iteration number 5 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	5.186	8	.737

Contingency Table for Hosmer and Lemeshow Test

Step		Apoyo a políticas públicas en Ciencia y Tecnología = No apoya		Apoyo a políticas públicas en Ciencia y Tecnología = Apoya		Total
		Observed	Expected	Observed	Expected	
		1	172	170.398	114	
	2	97	91.636	189	194.364	286
	3	73	75.089	213	210.911	286
	4	71	65.426	215	220.574	286
	5	52	57.574	234	228.426	286
	6	42	50.403	244	235.597	286
	7	40	43.506	246	242.494	286
	8	36	36.278	250	249.722	286
	9	33	28.627	253	257.373	286
	10	20	17.062	263	265.938	283

Classification Table<sup>a</sup>

Observed		Predicted		Percentage Correct
		Apoyo a políticas públicas en Ciencia y Tecnología		
		No apoya	Apoya	
Step 1	Apoyo a políticas públicas en Ciencia y Tecnología	126	510	19.8
	No apoya	54	2167	97.6
	Overall Percentage			80.3

- a. The cut value is .500

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
							Lower	Upper
Step 1 <sup>a</sup>								
SEXO	.296	.100	8.788	1	.003	1.345	1.106	1.635
EDAD	-.001	.003	.179	1	.672	.999	.992	1.005
ESCOLARIDAD	.038	.013	9.168	1	.002	1.039	1.013	1.065
CYTenTV	.008	.004	4.960	1	.026	1.008	1.001	1.015
CYTRADIO	.001	.006	.040	1	.841	1.001	.990	1.013
CYTPERIODICO	.005	.004	1.116	1	.291	1.005	.996	1.013
CYTREVISTAS	.007	.003	5.367	1	.021	1.007	1.001	1.014
CALIFICACION	.012	.004	11.327	1	.001	1.012	1.005	1.019
S4P23_2	.069	.069	1.003	1	.317	1.072	.936	1.227
S4P23_6	.008	.055	.021	1	.885	1.008	.906	1.122
S4P25_3	-.074	.059	1.555	1	.212	.929	.827	1.043
S4P25_7	.155	.060	6.734	1	.009	1.168	1.039	1.313
S4P26_3	.142	.076	3.452	1	.063	1.152	.992	1.338
S4P26_6	-.069	.086	.654	1	.419	.933	.789	1.104
S4P26_8	.084	.065	1.662	1	.197	1.087	.957	1.234
S4P26_10	.171	.073	5.420	1	.020	1.186	1.027	1.370
S4P26_11	.002	.071	.000	1	.982	1.002	.872	1.151
S4P28_4	.141	.058	5.980	1	.014	1.151	1.028	1.289
S4P33_2_1	.041	.066	.394	1	.530	1.042	.916	1.187
S4P33_2_2	.009	.071	.018	1	.894	1.010	.878	1.161
S4P33_2_3	-.051	.065	.609	1	.435	.950	.837	1.080
S4P33_2_4	.193	.057	11.265	1	.001	1.213	1.083	1.357
S4P33_2_5	.017	.054	.095	1	.758	1.017	.915	1.130
S4P33_2_6	-.017	.061	.080	1	.777	.983	.871	1.109
S4P33_2_7	.093	.063	2.159	1	.142	1.097	.969	1.242
Constant	-2.567	.378	46.241	1	.000	.077		

a. Variable(s) entered on step 1: SEXO, EDAD, ESCOLARIDAD, CYTenTV, CYTRADIO, CYTPERIODICO, CYTREVISTAS, CALIFICACION, S4P23\_2, S4P23\_6, S4P25\_3, S4P25\_7, S4P26\_3, S4P26\_6, S4P26\_8, S4P26\_10, S4P26\_11, S4P28\_4, S4P33\_2\_1, S4P33\_2\_2, S4P33\_2\_3, S4P33\_2\_4, S4P33\_2\_5, S4P33\_2\_6, S4P33\_2\_7.

# ANEXO 2. EJERCICIO ESTADÍSTICO DEFINITIVO

## Logistic Regression

Notes	
Output Created	03-JUL-2018 15:00:48
Comments	
Input	F:\New Folder\BASE.sav
Data	DataSet2
Active Dataset	<none>
Filter	<none>
Weight	<none>
Split File	<none>
N of Rows in Working Data File	2858
Missing Value Handling	User-defined missing values are treated as missing
Syntax	LOGISTIC REGRESSION VARIABLES APOYOCYT /METHOD=ENTER Sexo Escolaridad Consulta de CyT en televisión Consulta de CyT en revistas Calificación Los científicos deben ser libres de llevar a cabo sus investigaciones a su antojo, siempre y cuando las hagan bajo reglas éticas. Debería haber más mujeres dedicadas a la investigación científica en nuestro país. La ciencia tiene
Resources	
Processor Time	00:00:00.06
Elapsed Time	00:00:00.08

[DataSet2] F:\New Folder\BASE.sav

Case Processing Summary		
Unweighted Cases <sup>a</sup>	N	Percent
Selected	2857	100.0
Included in Analysis		
Cases	1	.0
Missing Cases		
Total	2858	100.0
Unselected Cases	0	0.0
Total	2858	100.0

a. If weight is in effect, see classification table for the total number of cases.

### Dependent Variable Encoding

Original Value	Internal Value
No apoya	0
Apoya	1

### Block 0: Beginning Block

Iteration History <sup>a,b,c</sup>		
Iteration	-2 Log likelihood	Coefficients Constant
Step 0	3039.602	1.110
1	3029.535	1.245
2	3029.523	1.251
3	3029.523	1.251
4	3029.523	1.251

a. Constant is included in the model.

b. Initial -2 Log Likelihood: 3029.523

c. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.

Observed		Predicted			
		Apoyo a políticas públicas	Apoya	Percentage Correct	
Step 0	Apoyo a políticas públicas en Ciencia y Tecnología	No apoya	0	636	0.0
		Apoya	0	2221	100.0
Overall Percentage					77.7

a. Constant is included in the model.

b. The cut value is .500

Variables in the Equation

Step	Model	B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	1.251	.045	773.166	1	.000	3.492

Variables not in the Equation

Step	Model	Score	df	Sig.
Step 0	Sexo	17.644	1	.000
	Escolaridad	151.580	1	.000
	Consulta de CyT en televisión	55.196	1	.000
	Consulta de CyT en revistas	43.234	1	.000
	Calificación	190.679	1	.000
	Los científicos deben ser libres de llevar a cabo sus investigaciones a su antojo, siempre y cuando las hagan bajo reglas éticas.	104.160	1	.000
	Debería haber más mujeres dedicadas a la investigación científica en nuestro país.	116.787	1	.000
	La ciencia tiene una imagen muy negativa para la sociedad.	87.710	1	.000
	El desarrollo tecnológico origina una manera de vivir artificial y deshumanizada.	104.884	1	.000
Overall Statistics		362.529	9	.000

Block 1: Method = Enter

Iteration	Iteration History <sup>a,b,c,d</sup>										
	-2 Log likelihood	Constant	Sexo	Escolaridad	Consulta de CyT en televisión	Consulta de CyT en revistas	Calificación	Los científicos deben ser libres de llevar a cabo sus investigaciones a su antojo, siempre y cuando las hagan bajo reglas éticas.	Debería haber más mujeres dedicadas a la investigación científica en nuestro país.	La ciencia tiene una imagen muy negativa para la sociedad.	El desarrollo tecnológico origina una manera de vivir artificial y deshumanizada.
Step 1	2738.461	-1.689	.183	.034	.003	.003	.010	.132	.183	.137	.165
2	2685.433	-2.354	.291	.045	.007	.006	.014	.171	.232	.168	.219
3	2683.725	-2.472	.315	.046	.009	.008	.015	.178	.240	.172	.216
4	2683.719	-2.475	.316	.046	.009	.008	.015	.179	.241	.172	.216
5	2683.719	-2.475	.316	.046	.009	.008	.015	.179	.241	.172	.216

a. Method: Enter

b. Constant is included in the model.

c. Initial -2 Log Likelihood: 3029.523

d. Estimation terminated at iteration number 5 because parameter estimates changed by less than .001.

Omnibus Tests of Model Coefficients			
Step		Chi-square	Sig.
Step 1	Block	345,804	.000
	Model	345,804	.000

Model Summary			
Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	2683,719 <sup>a</sup>	.114	.174

a. Estimation terminated at iteration number 5 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test			
Step	Chi-square	df	Sig.
1	6,893	8	.548

Contingency Table for Hosmer and Lemeshow Test						
Step 1		en Ciencia y Tecnología =		en Ciencia y Tecnología =		Total
		Observed	Expected	Observed	Expected	
1	1	167	166,211	119	119,789	286
	2	91	93,572	195	192,428	286
	3	84	76,258	202	209,742	286
	4	69	65,921	217	220,079	286
	5	48	57,925	238	228,075	286
	6	49	50,622	237	235,378	286
	7	38	43,723	248	242,277	286
	8	38	36,381	248	249,619	286
	9	36	28,350	250	257,650	286
	10	16	17,037	267	265,963	283

Classification Table <sup>a</sup>					
Observed		Predicted	Apoyo a políticas públicas		Percentage Correct
			No apoya	Apoya	
Step 1	Apoyo a políticas públicas en Ciencia y Tecnología	No apoya	119	517	18,7
		Apoya	49	2172	97,8
	Overall Percentage				80,2

a. The cut value is .500

Variables in the Equation								95% C.I. for EXP(B)	
Step 1 <sup>a</sup>		B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
	Sexo	.316	.099	10,199	1	.001	1,371	1,130	1,664
	Escolaridad	.046	.012	15,051	1	.000	1,047	1,023	1,072
	Consulta de CyT en televisión	.009	.004	6,328	1	.012	1,009	1,002	1,016
	Consulta de CyT en revistas	.008	.003	6,508	1	.011	1,008	1,002	1,014
	Calificación	.015	.003	18,165	1	.000	1,015	1,008	1,021
	Los científicos deben ser libres de llevar a cabo sus investigaciones a su antojo, siempre y cuando las hagan bajo reglas éticas.	.179	.056	10,225	1	.001	1,196	1,072	1,334
	Debería haber más mujeres dedicadas a la investigación científica en nuestro país.	.241	.062	14,877	1	.000	1,272	1,126	1,438
	La ciencia tiene una imagen muy negativa para la sociedad.	.172	.055	9,933	1	.002	1,188	1,067	1,322
	El desarrollo tecnológico origina una manera de vivir artificial y deshumanizada.	.216	.052	17,228	1	.000	1,241	1,121	1,374
	Constant	-2,475	.254	94,774	1	.000	.084		

a. Variable(s) entered on step 1: Sexo, Escolaridad, Consulta de CyT en televisión, Consulta de CyT en revistas, Calificación, Los científicos deben ser libres de llevar a cabo sus investigaciones a su antojo, siempre y cuando las hagan bajo reglas éticas.

Correlation Matrix										
Step 1	Constant	Sexo	Escolaridad	Consulta de CyT en televisión	Consulta de CyT en revistas	Calificación	Los científicos deben ser libres de llevar a cabo sus investigaciones a su antojo, siempre y cuando las hagan bajo reglas éticas.	Debería haber más mujeres dedicadas a la investigación científica en nuestro país.	La ciencia tiene una imagen muy negativa para la sociedad.	El desarrollo tecnológico origina una manera de vivir artificial y deshumanizada.
	1,000									
		1,000								
			1,000							
				1,000						
					1,000					
						1,000				
							1,000			
								1,000		
									1,000	
										1,000

## BIBLIOGRAFÍA

- [1] **Agenda Ciudadana de Ciencia, Tecnología e Innovación** (2012). <http://www.agendaciudadana.mx/index.php/2015-07-10-21-33-23/que-es-agenda-ciudadana>.
- [2] **BARRIO ALONSO, Cipriano** (2008): La apropiación social de la ciencia: nuevas formas. *Revista iberoamericana de ciencia, tecnología y sociedad* [online]. 2008, vol.4, n.10 [citado 2017-02-08], pp. 213-225. ISSN 1850-0013.
- [3] **CONACYT**. Informe general del estado de la ciencia y la tecnología, 1999. México.
- [5] **CONACYT**. Informe general del estado de la ciencia y la tecnología, 2016. México.
- [5] **CONACYT-INEGI** (2013): Encuesta sobre la percepción pública de la ciencia y la tecnología en México 2013. Síntesis Metodológica. México.
- [6] **DB-City**: Clasificación años promedio de escolarización (2010). Consulta en <https://es.db-city.com/Pa%C3%ADs--A%C3%B1os-promedio-de-escolarizaci%C3%B3n>. Fecha de última consulta: 23/09/2018.
- [7] **European Commission** (2010). Special Eurobarometer 340: Science and technology report. Brussels, Belgium.
- [8] **IMD World Competitiveness Center** (2016). *IMD World competitiveness yearbook 2016* Institute for Management Development, Lausanne, Switzerland. ISSN 1026-2628.
- [9] **Instituto de Estadística de la UNESCO** (2016): Indicadores de educación. UNESCO, Paris, Francia. Data extracted on 30 Aug 2018 20:44 UTC (GMT) from UIS.Stat.
- [10] **INEGI** (2018). Indicadores de nivel educativo. Página electrónica Cuéntame de México. <http://cuentame.inegi.org.mx/> . Última consulta 26-09-2018
- [11] **JUDGE, George G.**; Carter Hill R.; Griffiths, William E.; Lutkepohl, Helmut y Lee, Tsoung-Chao (1988): *Introduction to the theory and practice of econometrics*, segunda edición. John Wiley and Sons.
- [12] **LÓPEZ CERREZO, J. A.** y M. Cámara Hurtado (2005): *Apropiación social de la ciencia*, en *Percepción social de la ciencia y la tecnología*. España 2004, Madrid, FECYT.
- [13] **MILLER, John D.** (1996): *Public understanding of science and technology in OECD countries: A comparative analysis*. Symposium on public understanding of science and technology. Tokyo 5-6 nov. 1996. OECD.
- [14] **National Science Board** (2014). *Science and Engineering Indicators*, Chapter 7. Arlington VA: National Science Foundation (NSB 14-01).
- [15] **OECD** (2015), *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities*, OECD Publishing, Paris, Francia.
- [16] **OECD** (2016), *PISA 2015 Results (Volume I): Excellence and equity in education*, PISA, OECD Publishing, Paris, Francia.

- [17] **Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura** (1978): Recomendación relativa a la normalización internacional de las estadísticas de ciencia y tecnología. UNESCO, Paris, Francia.
- [18] **PEÑA, Daniel** (2002): Análisis de datos multivariantes. McGraw-Hill Interamericana de España, Madrid.
- [19] **POLINO, Carmelo**; Fazio, María Eugenia y Vaccarezza, Leonardo (2003): Medir la percepción pública de la ciencia en los países iberoamericanos. Aproximación a problemas conceptuales. Revista Iberoamericana de Ciencia, Tecnología, Sociedad e Innovación. OEI. Número 5 / Enero - Abril 2003.
- [20] **QS Top universities** (2018). QS World university rankings. Consulta en <https://www.topuniversities.com/>. Última consulta: 23/09/2018.
- [21] **Real Academia Española**. (2001). Diccionario de la lengua española (22nd ed.). Madrid, Spain.
- [22] **Red Iberoamericana de Indicadores de Ciencia y Tecnología, RICYT** (2015). Manual de Antigua: Indicadores de percepción pública de la ciencia y la tecnología. Antigua Guatemala.
- [23] **Sánchez V., Francisco** (2017). Regresión Logística. Diplomado de Estadística, ITAM, México.
- [24] **SNIEG**. Ley del Sistema Nacional de Información Estadística y Geográfica. Última reforma DOF 25-06-2018.