

TRANSICIÓN DE

ESTADÍSTICA

A

CIENCIA DE DATOS

EDITORES LITERARIOS

JULIA AURORA MONTANO RIVAS

CECILIA CRUZ LÓPEZ

ERNESTO PEDRO MENÉNDEZ ACUÑA

EMMANUEL MORALES GARCÍA

ISBN: 978-607-5893-50-1



TRANSICIÓN DE ESTADÍSTICA A CIENCIA DE DATOS

EDITORES LITERARIOS

Julia Aurora Montano Rivas, Cecilia Cruz López, Ernesto Pedro Menéndez Acuña, Emmanuel Morales García

AUTORES

Amy Yamilette Loeza Beureth, Angel Fernando Argüello Ortiz, Blandina, Bernal Morales, Blanca Xochilt Muñoz Vargas, Bulmaro Juárez Hernández, Cecilia Cruz López, Emmanuel Morales García, Ernesto Pedro Menéndez Acuña, Hortensia Josefina Reyes Cervantes, Iván Ruiz Hernández, Jeysira Jacqueline Dorantes Carrión, Jesús Hernández Suárez, José Juan Muñoz León, Juan Andrés Sánchez García, Juan Ruiz Ramírez, Julia Aurora Montano Rivas, Leonardo Ailines Genis, Luis Enrique Gómez Medina, María Teresa Leal Ascencio, Mario Eduardo Leal López, Martín Alfonso López Ramírez, Myrna Enedelia González Meneses, Noé Velázquez Rosas, Rebeca Menchaca García, Roberto Gallardo del Ángel, Guadalupe Sánchez Sánchez, Sergio Francisco Juárez Cerrillo, Sulayka Florencia Castelán Culebro, María de Lourdes Velasco Vázquez, Yesid Carvajal Escobar, Citlali Guadalupe Lozada Hernández, Edson Valdés Iglesias, Alexis Iván Mora Argüello, José Luis Hernández González, Christopher Cernichiario Reyna.

©RED IBEROAMERICANA DE ACADEMIAS DE INVESTIGACIÓN A.C. 2025



EDITA: RED IBEROAMERICANA DE ACADEMIAS DE INVESTIGACIÓN A.C.
DUBLÍN 34, FRACCIONAMIENTO MONTE MAGNO
C.P. 91190. XALAPA, VERACRUZ, MÉXICO.
CEL 2282386072
www.redibai.org
redibai@hotmail.com

ISBN: 978-607-5893-50-1



Sello editorial: Red Iberoamericana de Academias de Investigación, A.C.
(978-607-5893)
Primera Edición, Xalapa, Veracruz, México.
No. de ejemplares: 2
Presentación en medio electrónico digital
Formato PDF 10 MB
Fecha de aparición 28/11/2025
ISBN 978-607-5893-50-1

DICTAMEN EDITORIAL

La presente obra fue arbitrada y dictaminada en dos procesos; el primero, fue realizado por el **COMITÉ EDITORIAL RED IBEROAMERICANA DE ACADEMIAS DE INVESTIGACIÓN A.C.** con sede en México; que sometió a los capítulos incluidos en la obra a un proceso de dictaminación a doble ciego para constatar de forma exhaustiva la temática, pertinencia y calidad de los textos en relación a los fines y criterios académicos de la misma, cumpliendo así con la primera etapa del proceso editorial. El segundo proceso de dictaminación estuvo a cargo del **COMITÉ CIENTÍFICO RED IBEROAMERICANA DE ACADEMIAS DE INVESTIGACIÓN A.C.**; donde se seleccionaron expertos en el tema para la evaluación de los capítulos de la obra y se procedió con el sistema de dictaminación a doble ciego. Cabe señalar que previo al envío a los dictaminadores, todo trabajo fue sometido a una prueba de detección de plagio. Una vez concluido el arbitraje de forma ética y responsable y por acuerdo del Comité Editorial y Científico de la Red Iberoamericana de Academias de Investigación A.C. (REDIBAI), se dictamina que la obra **“TRANSICIÓN DE ESTADÍSTICA A CIENCIA DE DATOS”** cumple con la relevancia y originalidad temática, la contribución teórica y aportación científica, rigurosidad y calidad metodológica, actualidad de las fuentes que emplea, redacción, ortografía y calidad expositiva.

Dr. Daniel Armando Olivera Gómez

Director Editorial

Sello Editorial: Red Iberoamericana de Academias de Investigación, A.C.

(978-607-5893)

Dublín 34, Residencial Monte Magno

C.P. 91190. Xalapa, Veracruz, México.

Cel 2282386072

CERTIFICACIÓN EDITORIAL

RED IBEROAMERICANA DE ACADEMIAS DE INVESTIGACIÓN A.C. (REDIBAI) con sello editorial **No. 978-607-5893** otorgado por la Agencia Mexicana de ISBN, hace constar que el libro **“TRANSICIÓN DE ESTADÍSTICA A CIENCIA DE DATOS”** registrado con el **ISBN 978-607-5893-50-1** fue publicado por nuestro sello editorial con fecha de aparición del 28 de noviembre de 2025 cumpliendo con todos los requisitos de calidad científica y normalización que exige nuestra política editorial.

Fue evaluado por pares académicos externos y aprobado por nuestro Comité Editorial y Científico.

Todos los soportes concernientes a los procesos editoriales y de evaluación se encuentran bajo el poder Editorial de **RED IBEROAMERICANA DE ACADEMIAS DE INVESTIGACIÓN A.C. (REDIBAI)**, los cuales están a disposición de la comunidad académica interna y externa en el momento que se requieran.

La normativa editorial y repositorio se encuentran disponibles en la página **<http://www.redibai-myd.org>**

Doy fe.

Dr. Daniel Armando Olivera Gómez

Director Editorial

Sello Editorial: Red Iberoamericana de Academias de Investigación, A.C.
(978-607-5893)

Dublín 34, Residencial Monte Magno

C.P. 91190. Xalapa, Veracruz, México.

Cel 2282386072

ÍNDICE

PROLOGO.....	1
SEGMENTACIÓN DE MUNICIPIOS DE VERACRUZ SEGÚN LA INCIDENCIA DEL VIH UTILIZANDO ANÁLISIS CLUSTERING.....	2
ESTIMACIÓN DE LA OFERTA LABORAL INCORPORANDO MÉTODOS NO-PARAMÉTRICOS.....	14
ANÁLISIS DE TENDENCIAS DE LA CONCENTRACIÓN DEL OXÍGENO DISUELTO Y DE LA TEMPERATURA DE LA REGIÓN HIDROLÓGICA GOLFO CENTRO, MÉXICO	32
CORRELACIÓN ESPACIAL DE ACTIVIDADES ECONÓMICAS DE XALAPA BASADO EN EL ANÁLISIS CLÚSTER Y DE VALOR ATÍPICO (I ANSELIN LOCAL DE MORAN)	42
APLICACIÓN DE SOFTWARE LIBRE EN LA MINERÍA DE DATOS: R Y PYTHON COMO HERRAMIENTAS DE ANÁLISIS.....	61
ÍNDICE NACIONAL DE PRECARIEDAD LABORAL: TRATAMIENTO ESTADÍSTICO DE BASES DE DATOS (ENOE)	77
SERVICIOS ECOSISTÉMICOS EN EL SECTOR DE USO DE SUELO: ¿EXISTE EVIDENCIA DE SU CONTRIBUCIÓN AL BIENESTAR?	89

ÍNDICE

¿HIJOS PORELECCIÓN? LA SITUACIÓN DE LA FECUNDIDAD EN MÉXICO, 2014.....	101
--	-----

TIPOS DE VIOLENCIA QUE EXPERIMENTAN LOS ESTUDIANTES UNIVERSITARIOS. UN ANÁLISIS ESTADÍSTICO CON (SPSS).....	119
---	-----

ANÁLISIS ESTADÍSTICO DE LA RELACIÓN ENTRE LA SINTOMATOLOGÍA DE ANSIEDAD Y DEPRESIÓN CON EL RENDIMIENTO ESCOLAR EN UNA MUESTRA DE ALUMNOS BACHILLERES.....	136
---	-----

EVALUACIÓN DE LA MICROPROPAGACIÓN IN VITRO DE HÍBRIDOS SELECTOS DEL GÉNERO VANILLA USANDO UN DISEÑO COMPLETAMENTE AL AZAR	143
---	-----

REPRESENTACIONES SOCIALES DE ACADÉMICOS DE LA UV SOBRE LA VIOLENCIA. NUBE DE PALABRAS Y ANÁLISIS DE SIMILITUD-SOFTWARE IRAMUTEQ.....	152
--	-----

EFFECTIVIDAD DEL MECANISMO DE TRANSMISIÓN DE LA POLÍTICA MONETARIA EN MÉXICO (2020-2023)	165
--	-----

ANÁLISIS DE LA CANTIDAD DE ACCIDENTES DE TRÁNSITO MENSUALES EN EL ESTADO DE PUEBLA MEDIANTE EL MODELO NORMAL UNIVARIADO CON PUNTOS DE CAMBIO	185
--	-----

Prólogo

Durante siglos, la Estadística ha sido el lenguaje silencioso detrás de los grandes descubrimientos científicos, las políticas públicas y las decisiones estratégicas que moldean nuestras sociedades. Nació como una herramienta para describir el mundo y se transformó en una disciplina capaz de explicarlo y predecirlo. Sin embargo, el siglo XXI trajo consigo un cambio de escala, de velocidad y de complejidad: los datos se volvieron ubicuos, crecieron en volumen, variedad y velocidad, y con ellos emergió una nueva forma de mirar la realidad: la Ciencia de Datos.

Este libro, *Transición de Estadística a Ciencia de Datos*, busca tender un puente entre dos mundos que, aunque distintos en apariencia, comparten un mismo corazón: la búsqueda del conocimiento a partir de la evidencia. La Estadística provee el rigor, la inferencia y la interpretación; la Ciencia de Datos suma la automatización, la programación y la potencia del cómputo. Juntas, permiten transformar datos en decisiones y decisiones en impacto social.

La transición no implica ruptura, sino evolución. Así como la Estadística nació de la necesidad de ordenar la información del Estado, la Ciencia de Datos surge para comprender el universo digital en el que vivimos. Hoy, los algoritmos no sustituyen al pensamiento crítico del estadístico; lo potencian. La programación no elimina la reflexión metodológica; la expande. Y la inteligencia artificial no anula la intuición humana; la complementa.

Esta obra invita al lector —profesional, estudiante o investigador— a recorrer este proceso de integración conceptual y práctica. A comprender cómo los fundamentos de la Estadística sustentan los modelos de Machine Learning; cómo los principios del diseño experimental dialogan con los entornos de análisis masivo; y cómo el pensamiento estadístico sigue siendo el eje ético y epistemológico de la Ciencia de Datos.

En última instancia, este libro celebra una idea esencial: que la verdadera transición no es tecnológica, sino cultural. La Ciencia de Datos no sustituye a la Estadística; la honra, la renueva y la proyecta hacia los desafíos del futuro.

Editores

SEGMENTACIÓN DE MUNICIPIOS DE VERACRUZ SEGÚN LA INCIDENCIA DEL VIH UTILIZANDO ANÁLISIS CLUSTERING

Citlali Guadalupe Lozada Hernández, Cecilia Cruz López, Emmanuel Morales García,

Luis Enrique Gómez Medina

Facultad de Estadística e Informática

Universidad Veracruzana

e-mail: gS21013278@egresados.uv.mx, ceccruz@uv.mx, emmorales@uv.mx, luisgomez04@uv.mx

RESUMEN

El presente trabajo tiene como objetivo segmentar los municipios de Veracruz con casos activos de VIH en distintos clústeres, considerando la prevalencia de la enfermedad (porcentaje de casos totales y activos sobre la población), el perfil demográfico (distribución por sexo y edad promedio) y las condiciones de rezago social (índice o grado de rezago). A diferencia de los análisis tradicionales, esta técnica permite identificar patrones que podrían pasar desapercibidos. Los resultados muestran tres grupos epidemiológicos con características diferenciadas, lo que evidencia que una estrategia única no sería eficaz. La segmentación territorial facilita diseñar acciones específicas de prevención, detección y atención del VIH. El análisis clúster resulta especialmente útil en contextos epidemiológicos, ya que adapta las intervenciones a dinámicas locales cambiantes.

Palabras claves: Clasificación no supervisada, georreferencias, análisis multivariado, salud

INTRODUCCIÓN

Desde la década de los ochenta el VIH (Virus de la Inmunodeficiencia Humana) es un problema de salud pública a nivel global, desde que se presentaron los primeros casos en México en 1983, estos han ido en aumento, provocando así una crisis sanitaria tanto por la falta de medicación como por las enfermedades derivadas del SIDA.

Determinados estados en el país presentan una mayor cantidad de casos, por ejemplo, a raíz del inicio de la epidemia en territorio nacional, generalmente la ciudad de México es la entidad federativa con más casos diagnosticados en el país año con año (Centro Nacional para la Prevención y Control del VIH y el SIDA, 2023)., esto no solo se debe a aspectos epidemiológicos, sino también a diversos factores sociodemográficos o económicos, tales como la prevalencia de la enfermedad, porcentaje de casos

dependiendo el sexo, la edad promedio de los individuos y el índice de rezago social, así como al seguimiento de patrones que incrementan las infecciones. De igual forma Estado de México y Veracruz son municipios con gran carga de infecciones en términos absolutos, posicionándose en segundo y tercer lugar respectivamente. La elección de hacer uso de números absolutos radica en el interés por identificar las entidades con mayor carga total del sistema de salud, independientemente de que tamaño tenga la población, lo anterior es de utilidad para observar en que estados se concentra la mayor cantidad de personas afectadas.

En el presente trabajo se aplican técnicas de análisis *clustering* (específicamente *K-means*) para segmentar los municipios con casos activos de VIH en función de las variables mencionadas con anterioridad, para identificar grupos que compartan características epidemiológicas ya que puede permitir revelar patrones que en análisis tradicionales pasan desapercibidos. Cabe mencionar que la enfermedad de VIH/SIDA no corresponden a una epidemia generalizada, más bien se trata de una epidemia concentrada en ciertos grupos.

El lograr identificar distintos grupos y las similitudes que tienen entre sí ayuda a entender como múltiples factores y prácticas influyen en el aumento y control de casos de VIH en los municipios del estado de Veracruz. El enfoque de dicho trabajo puede ser útil para que los responsables de las políticas públicas de salud prioricen acciones para determinados municipios y no se den soluciones generalizadas que podrían no ser aplicables en todos los contextos.

Antecedentes

Oster et al. (2018) analizan 156,553 secuencias parciales de la polimerasa del VIH-1 registradas en el Sistema Nacional de Vigilancia del VIH, e infieren los clústeres de transmisión utilizando 0.5% y 1.5% como umbrales de distancia genética, así como dos períodos de tiempo para los diagnósticos: todos los años disponibles y 2013-2015 (sólo diagnósticos recientes). Se busca detectar conglomerados con transmisión rápida del VIH, lo cual es fundamental para que los esfuerzos se centren en aquellas poblaciones que necesitan intervenciones de prevención más intensiva, para lograrlo, se toma en cuenta únicamente a aquellos conglomerados con al menos 5 personas diagnosticadas en el año 2015 y se hace uso de un análisis llamado “reloj molecular” que permite estimar cuándo surgió el virus común del que surgieron todas las infecciones dentro del grupo. Como resultados se encontró que utilizando un umbral de distancia de 1.5% se identificaron 103 conglomerados de rápido crecimiento haciendo uso de todos los diagnósticos disponibles y 73 conglomerados haciendo uso únicamente de los diagnósticos recientes. Por

su parte, un umbral de 0.5% detectó 15 conglomerados con todos los diagnósticos disponibles y 13 con los diagnósticos recientes.

Pérez Parra et. al (2015) estudiaron a 693 pacientes diagnosticados con VIH-1 durante los años de 2005-2012 en Andalucía oriental, en los hospitales de Granada, Almería y Jaén. De los 693 pacientes se disponía de una secuencia fasta del gen pol (posiciones 4-99 de la proteasa y 38-247 de la transcriptasa reversa), las cuales fueron obtenidas mediante el ensayo Trugene HIV-1 Genotyping Kit (Siemens [NAD]). 298 de los 693 pacientes se asociaron en 77 clústeres distintos. La mayoría de los conglomerados se conformaron por parejas ($n = 49$) de hombres que tienen relaciones sexuales con hombres ($n = 26$), de nacionalidad española ($n = 37$). 48% de los grupos estaba conformado por nativos españoles, 23% incluía a algún extranjero en sus integrantes y solo 3% se componía únicamente por extranjeros. 73.5% de los pacientes tienen menos de 45 años y 31.6% viven en el área metropolitana de Granada. La transmisión se produce más frecuentemente (66%) entre pacientes infectados con subtipos B y subtipos no B (22%), mientras que el resto (12%) se produjo en pacientes infectados con distinto subtipo (B y recombinante B-D).

Elizondo et. al (2014) analizan la percepción y actitud de las personas seropositivas respecto a los servicios odontológicos. Se contó con la participación de 134 voluntarios (70% hombres y 30% mujeres) de Nuevo León, México, en 2014, los cuales respondieron un cuestionario estructurado de tipo analítico, auto-administrado y anónimo. Mediante clústeres no jerárquico se analizaron variables socio demográficas, así como la percepción sobre los servicios y prestadores de servicios odontológicos tanto públicos como privados; se exploró la percepción del estigma asociado al VIH/SIDA por medio de la escala tipo Likert. Como resultados se identificaron desigualdades en la búsqueda de atención odontológica ya sea pública o privada. el análisis clúster logró identificar 3 grupos: usuarios que no han experimentado estigma ni discriminación; usuarios que no han experimentado estigma ni discriminación, pero sienten un poco de preocupación y usuarios que han experimentado estigma, discriminación y sienten preocupación, con 85.0%, 12.7% y 2.3% respectivamente. En conclusión, el porcentaje de estigma y discriminación es bajo, empero, la mayoría de los pacientes no comunican su diagnóstico por temor a sufrir rechazo.

Moral (2007) busca describir conductas de sexo seguro, protegido y de riesgo; uso de métodos anticonceptivos e identificar grupos de sujetos en vínculo con el uso de preservativos en estudiantes universitarios para prevenir la infección por VIH. Se utilizó una muestra de 395 estudiantes (que recibían capacitación para ser instructores de educación sexual) de educación superior en una universidad mexicana entre agosto de 2004 y agosto de 2005, los datos se analizaron con las técnicas de regresión logística, análisis de correspondencia múltiple y de clúster por el método de Ward. El dendograma

identificó tres clústeres: el primero, con 43 sujetos, se caracteriza por no haber usado anticonceptivos en la primera relación (58%) o por recurrir al coito interrumpido (42%), siendo el preservativo el método más común (67%); se denomina “activo sexualmente, de riesgo de embarazo y contagio de VIH”. El segundo clúster, integrado por 39 sujetos, usó preservativo en la primera relación y lo mantiene como método principal (95%), con baja frecuencia sexual (82% pocas veces y 18% solo una vez), denominado “poco activo sexualmente, de bajo riesgo”. El tercero, con 50 sujetos, utilizó mayoritariamente preservativo en la primera relación (80%), aunque algunos no emplearon ningún método (12%), optaron por otro (6%) o coito interrumpido (2%); actualmente el preservativo es el más usado (66%), seguido de otros métodos (14%), píldora (12%), coito interrumpido (6%) y ninguno (2%), siendo el grupo con mayor actividad sexual y de “bajo riesgo de embarazo”.

MATERIALES Y MÉTODOS

Área de estudio

El estado de Veracruz es una de las treinta y dos entidades federativas que conforman México siendo la ciudad de Xalapa su capital (Figura 1). Cuenta con 212 municipios divididos en 10 regiones (Huasteca alta, Huasteca baja, Totonaca, Nautla, Capital, Las Montañas, Sotavento, Papaloapan, Los Tuxtlas, Olmeca) [SEFIPLAN 2005]. Según INEGI en el año 2020 la población de todo el estado era de aproximadamente 8,062,579 habitantes, representando el 6.4% de la población de todo el país, ocupando el cuarto lugar a nivel nacional con mayor número de habitantes, siendo Veracruz el municipio más poblado del estado, con una población de 607,209 habitantes, representa el 3.7% de la superficie del país con una extensión de 71,823.5 km² y una densidad de 112 personas por kilómetro cuadrado (INEGI, s.f.).

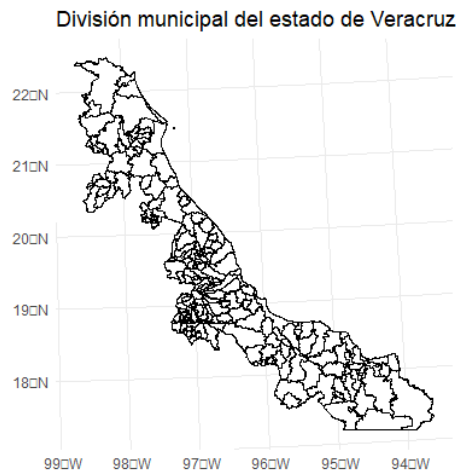


Figura 1. Mapa de Veracruz

Datos

Los datos utilizados para realizar el análisis se obtuvieron de la Secretaria de Salud y Servicios de Salud Veracruz (SESVR), contiene observaciones que abarcan del año 2020 al año 2023, contiene las siguientes variables (Tabla 1).

Tabla 1. Información de las variables

Variable	Código	Descripción	Tipo de variable
Clave municipio	cve_mun	Número con el que se identifica a cada uno de los municipios	Cualitativa
Municipio	Municipio	Nombre del municipio	Cualitativa
Longitud	Longitud	Valor de longitud de cada municipio	Cuantitativa
Latitud	Latitud	Valor de latitud de cada municipio	Cuantitativa
Casos positivos activos	Casos positivos activos	Número de casos de VIH que son tratados y permanecen activos en algún CAPASIT	Cuantitativa
Casos positivos totales	Casos positivos totales	Número de casos de VIH totales, tengan baja, suspensión o fallecimiento	Cuantitativa
Porcentaje de casos totales	Porcentaje_casostotales	Porcentaje de casos totales de VIH, tengan baja, suspensión o fallecimiento	Cuantitativa
Porcentaje de casos activos	Porcentaje_casosactivos	Porcentaje de casos de VIH que son tratados y permanecen activos en algún CAPASIT	Cuantitativa
Índice de rezago	Indice_rezago	Índice de rezago social por municipio (tomado del CONEVAL)	Cuantitativa

TRANSICIÓN DE ESTADÍSTICA A CIENCIA DE DATOS

Grado de rezago	Grado_rezago	Grado de rezago de cada uno de los municipios (muy alto, alto, medio, bajo, muy bajo)	Cuantitativa
Mortalidad	Mortalidad	Personas fallecidas por VIH en cada uno de los municipios	Cuantitativa
Porcentaje de mortalidad	porcentaje_mortalidad	Porcentaje de personas fallecidas por VIH en cada uno de los municipios	Cuantitativa
Centro de salud	centro_salud	CAPACITS que atienden los casos de VIH en cada uno de los municipios	Cualitativa
Casos masculinos	masculino_casos	Casos de VIH de sexo masculino por municipio	Cuantitativa
Casos femeninos	femenino_casos	Casos de VIH de sexo femenino por municipio	Cuantitativa
Porcentaje casos masculino	mas_porc	Porcentaje de casos de VIH en personas de sexo masculino por municipio	Cuantitativa
Porcentaje casos femenino	fem_porc	Porcentaje de casos de VIH en personas de sexo femenino por municipio	Cuantitativa
Promedio edad	promedio_edad	Edad promedio en cada uno de los municipios	Cuantitativa

Análisis estadístico

Las técnicas utilizadas para tratar los datos utilizados en el presente trabajo constan, en primera instancia de un análisis descriptivo de los datos, mediante gráficos de barras, e histogramas se conoce la distribución de los datos.

Por su parte, el análisis final comienza obteniendo el número óptimo de clústeres haciendo uso del método del codo y de la silueta. Una vez obtenidos el número ideal de grupos se procede a formarlos mediante k medias (k -means), posteriormente se conocen sus estadísticas promedio por clúster. Se mapean los grupos obtenidos para ver cómo se distribuyen en el espacio (estado de Veracruz).

RESULTADOS

Análisis descriptivo

En la Figura 2 se observa la frecuencia del grado de rezago de los 212 municipios del estado de Veracruz. 92 municipios tienen un grado bajo de rezago mientras que únicamente 15 de ellos tienen un grado muy alto.

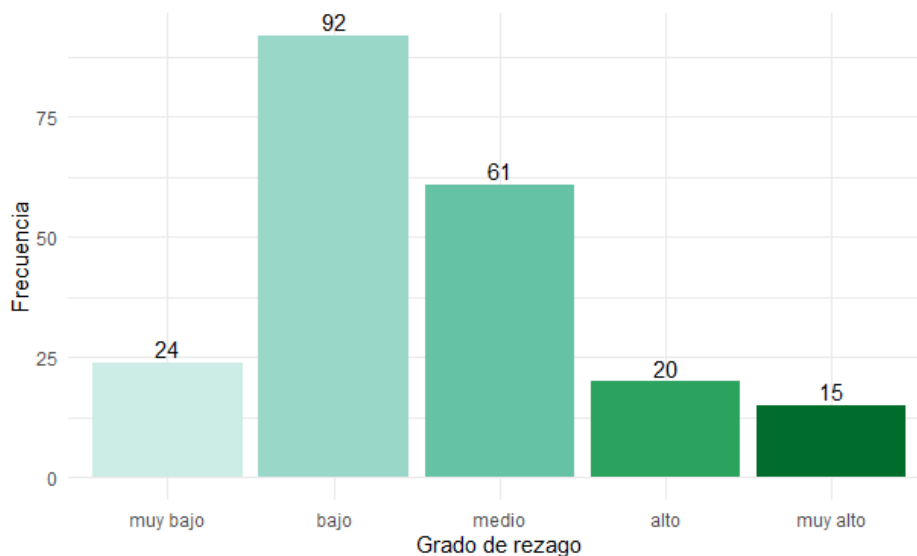


Figura 2. Frecuencia del rezago social

En la Figura 2 se observa que el género masculino presenta más de 4,500 casos de VIH mientras que hay cerca de 1,300 casos de personas del género femenino.

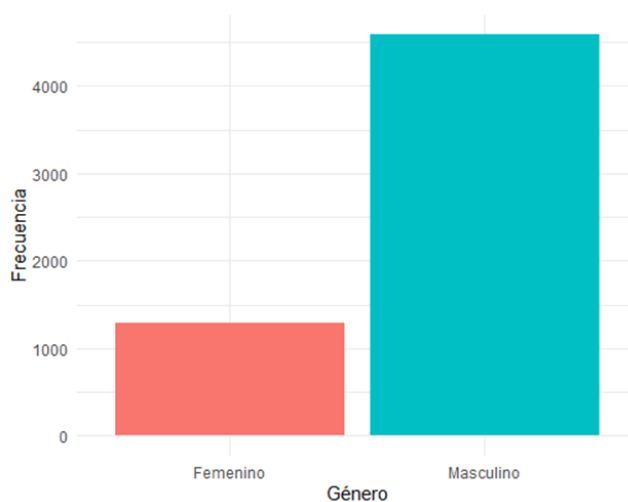


Figura 2. Frecuencia por género

En la Figura 3 se observa el promedio de edad en personas con VIH en Municipios del estado de Veracruz, se observa que aproximadamente de los 30 a los 40 años hay mayor concentración de casos, una población relativamente joven. De igual forma hay excepciones, siendo el caso del promedio de edad de 0 y 65 años.

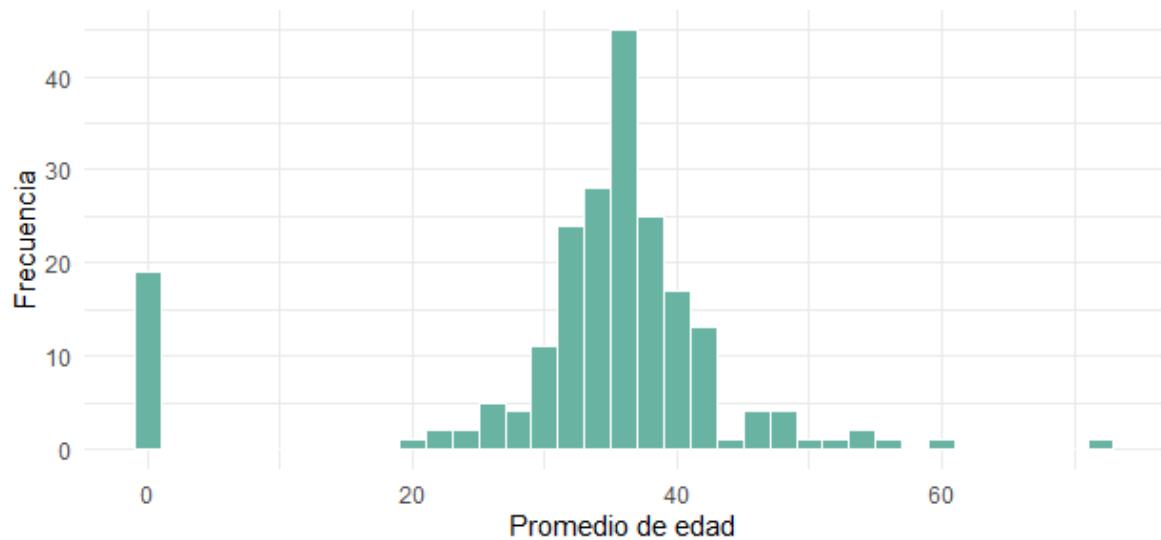


Figura 3. Distribución del Promedio de Edad en Municipios de Veracruz.

Análisis multivariante: Clúster k medias

Determinar el número óptimo de clústeres mediante el método de la silueta:

Como se observa en la Figura 4, el número óptimo de clústeres obtenido mediante el método de la silueta es 3. Mismo resultado obtenido mediante el método del codo, por lo que se considera que se deben crear 3 clústeres para agrupar las observaciones.

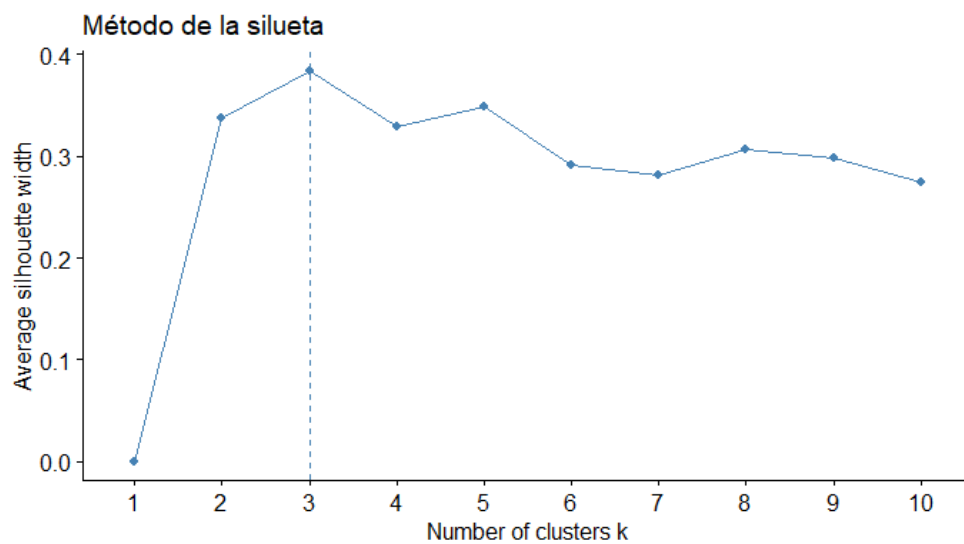


Figura 4. Número óptimo de clústeres mediante el método de la silueta.

Creación de clúster

El siguiente paso después de haber obtenido el número adecuado de grupos es proceder con la aplicar el algoritmo de agrupamiento k medias con dicho número de clústeres.

TRANSICIÓN DE ESTADÍSTICA A CIENCIA DE DATOS

En la Figura número 5 se observan los 3 grupos creados a partir de los datos obtenidos. Las dimensiones mostradas en ambos ejes representan una reducción dimensional con respecto a los datos originales, explicando un 81.4% de la varianza total, por su parte, los iconos representan los municipios (identificados por su número). Las observaciones ubicadas dentro de un mismo grupo son similares entre sí y diferentes en relación con el resto de los grupos. El clúster de color azul con iconos de círculo agrupa la mayoría de los municipios, el clúster de color amarillo con iconos de triángulo representa una agrupación más dispersa a la izquierda, por último, el clúster de color gris con iconos de cuadrado agrupa a la minoría de municipios, más definida arriba a la derecha.

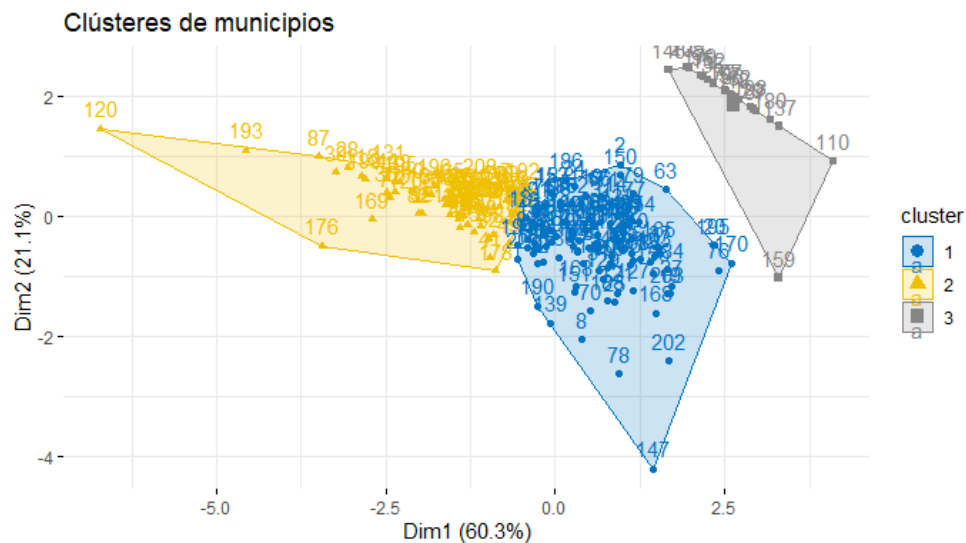


Figura 5. Clúster de municipios.

En la Tabla 2, se muestran las estadísticas promedio por clúster utilizando *k-means*, el clúster 1 cuenta con 114 municipios, el clúster 2 se conforma por 78 municipios, finalmente, el clúster 3 se conforma por 20 municipios.

Tabla 2. Estadísticas promedio por clúster.

Clúster	Media prevalencia	Media rezago	Media porcentaje casos masculino	Media edad promedio	Número de municipios
1	0.0193042	0.4910634	0.02170206	36.5986	114
2	0.05712335	-0.3902071	0.06381564	36.45654	78
3	0.00144232	0.9792603	0	1.3	20

El grupo 1 podría representar municipios con moderada prevalencia de VIH (1.93%), rezago social intermedio (49%), y una proporción baja de casos en hombres (2.17%), es una población adulta joven (36.6 años).

TRANSICIÓN DE ESTADÍSTICA A CIENCIA DE DATOS

El grupo 2 podría representar municipios con alta prevalencia de VIH (5.71%), bajo rezago social (zonas urbanas -0.39%), con mayoría de casos en hombres (6.38%), esto sugiere un patrón típico en contextos urbanos y concentraciones en hombres que tienen sexo con hombres, de igual forma es una población adulta joven (36.5 años).

El grupo 3 podría representar municipios altamente marginados (0.97%) con muy baja prevalencia reportada (0.14%), la edad promedio (1.3) está mal registrado.

Georreferencias del clúster k medias

En la Figura 5, se muestra el mapeo de los 3 clústeres obtenidos al aplicar k medias, el clúster 1 representa municipios marginados, con prevalencia moderada, donde el VIH afecta principalmente a mujeres jóvenes; el segundo clúster representa municipios urbanos, con alta incidencia de VIH, especialmente entre hombres adultos jóvenes (aquí se ubican la ciudad de Xalapa y Veracruz); el tercer clúster representa municipios con prevalencia intermedia, edad promedio más alta y pocos hombres entre los casos.



Figura 5. Mapeo de clúster en los municipios.

CONCLUSIONES

Se logró segmentar los municipios del estado de Veracruz en tres clústeres según la prevalencia de VIH, porcentaje de casos masculinos (sexo), promedio de edad e índice de rezago. En primera instancia fueron utilizados los 212 municipios del estado para crear los grupos, dando como resultado un clúster que conformado por municipios que no habían registrado casos activos de VIH, por lo que se prescindió de ellos y se contó con 183 municipios para realizar un segundo análisis.

Se caracterizaron los municipios mediante la prevalencia de VIH, perfil demográfico (porcentaje de casos por sexo y edad promedio) así como condiciones de rezago social (índice/grado de rezago social). Se obtuvieron las medidas de tendencia central, se encontró que son pocos los municipios en donde se

concentra la mayor cantidad de casos activos y totales de VIH, los cuales provocan que la media se eleve, de igual manera, existe una alta variabilidad debido a que hay municipios que no presentan ningún caso activo y otros que tienen hasta 689 casos activos.

En promedio hay 21.62 casos masculinos y 6.06 casos femeninos por municipio, existe un claro predominio masculino, pero con mayor dispersión en mujeres. La edad promedio se ubica en 33.22 años, con un mínimo de 0 y máximo de 73. De igual forma se realizaron dos visualizaciones, la primera muestra la distribución de casos activos de VIH en Veracruz y se observó que en la mayoría de los municipios no hay una gran concentración de casos activos, por su parte, en la segunda visualización se muestra el grado de rezago social y los casos activos de VIH, se observó que existen cinco niveles de rezago social (muy alto, alto, medio, bajo y muy bajo), la mayoría de los municipios tienen un rezago social bajo y medio con 0.15% de casos activos.

Referencias

- Centro Nacional para la Prevención y Control del VIH y el SIDA. (2023). *Sistema de vigilancia epidemiológica de VIH. Informe histórico de VIH 4to trimestre 2023*. Gobierno de México. <https://www.gob.mx/censida/documentos/epidemiologia-registro-nacional-de-casos-de-sida>
- Elizondo, J. E., Treviño, A. C., & Violant, D. (2015). La odontología y el estigma asociado al VIH. *Revista de saúde pública*, 49(79), 1–11. <https://doi.org/10.1590/S0034-8910.2015049005877>
- Instituto Nacional de Estadística y Geografía (INEGI). (s.f.). *Información por entidad federativa. En Cuéntame de México*. <https://cuentame.inegi.org.mx/monografias/informacion/ver/default.aspx?tema=me&e=30>
- Moral de la Rubia, J. (2007). Conducta sexual y uso del preservativo en estudiantes universitarios. *Medicina Universitaria*, 9(37), 173–180.
- Oster, A. M., France, A. M., Panneer, N., Bañez Ocfemia, M. C., Campbell, E., Dasgupta, S., Switzer, W. M., Wertheim, J. O., & Hernandez, A. L. (2018). Identifying clusters of recent and rapid HIV transmission through analysis of molecular surveillance data. *Journal of Acquired Immune Deficiency Syndromes*, 79(5), 543–550. <https://doi.org/10.1097/QAI.0000000000001856>

- Pérez-Parra, S., Chueca-Porcuna, N., Álvarez-Estevez, M., Mohamed, O., Collado, A., Vinuesa, D., Lozano, A. B., & García-García, F. (2015). Los estudios de resistencias a antirretrovirales como herramienta para el análisis de los clusters de transmisión del virus de la inmunodeficiencia humana. *Enfermedades Infecciosas y Microbiología Clínica*, 33(9), 603–608. <https://doi.org/10.1016/j.eimc.2014.11.016>
- Secretaría de Finanzas y Planeación de Veracruz, Comité de Planeación para el Desarrollo del Estado de Veracruz (SEFIPLAN-COPLADEVER). (2005). *Estudios regionales para la planeación. Información básica*.
- Secretaría de Salud de Veracruz. (2023). *Informe epidemiológico de VIH en Veracruz* 2023. Gobierno del Estado de Veracruz. <https://www.veracruz.gob.mx/salud/>

ESTIMACIÓN DE LA OFERTA LABORAL INCORPORANDO MÉTODOS NO-PARAMÉTRICOS

Roberto Gallardo Del Ángel, Julia Aurora Montano Rivas, Edson Valdés Iglesias.

Facultad de Economía, Facultad de Estadística e Informática

Universidad Veracruzana.

rogallardo@uv.mx, julmontano@uv.mx, edvaldes@uv.mx.

RESUMEN

El estudio analiza la oferta laboral con curvatura hacia atrás prevista por la teoría y las estadísticas oficiales, usando la ENOE 2020-T1. Para evitar sesgo de selección se aplicó la corrección de Heckman y se exploró la relación salario–horas mediante Lowess, diferenciando por hombres y mujeres. Se identifican dos tramos: uno creciente y elástico respecto al salario y otro decreciente e inelástico, con puntos de quiebre para ambos grupos. Al estimar la oferta laboral con salario, escolaridad e ingreso de otros miembros del hogar, se encontró que hombres con salarios altos tienen un incremento en su oferta laboral que es menos que proporcional al incremento al salario y es lo contrario si tienen salarios bajos. En las mujeres la respuesta es mayor, aunque las elasticidades son, en general, inelásticas, las de menor salario exhiben una oferta más elástica que las de mayor salario.

Palabras clave: Oferta laboral, Curva de oferta hacia atrás, Elasticidades salario–horas, Punto de quiebre, Corrección de Heckman, Lowess.

INTRODUCCIÓN

Existe una amplia literatura sobre la estimación de la oferta laboral en México. Entre ella, destaca el estudio de Gong y Van Soest (2002) que se centra en la oferta de trabajo de mujeres casadas en la Ciudad de México usando un modelo neoclásico estático con datos de la Encuesta Nacional de Empleo Urbano (ENEU) de 1992. Los autores estiman una elasticidad del salario dentro de la oferta de trabajo de 0.87 y una elasticidad del ingreso del cónyuge de -0.17, lo que implica que un aumento dentro de las ganancias del cónyuge reduce la oferta de trabajo femenina.

Gabriel Martínez (2012) estima las elasticidades de la reforma laboral en México utilizando los datos obtenidos de la ENOE de 2008 y 2009. Sus resultados muestran que las horas de trabajo no cambian significativamente tras una variación en su salario. También encuentra que a presencia de hijos aumenta

el trabajo de los hombres, pero lo reduce para las mujeres. Además, la afiliación a la seguridad social tiene un efecto positivo en las horas trabajadas específicamente para las mujeres.

Gordon H. Hanson (2005) analiza el impacto que tiene la emigración mexicana hacia Estados Unidos sobre las ofertas laborales de los mexicanos, con base en los censos de 1990 y 2000. Compara la evolución de los salarios entre entidades con alta y baja emigración histórica y encuentra que, en los estados más expuestos, la población masculina de 20 a 59 años se reduce alrededor de 10% respecto de otras entidades, lo que se asocia con incrementos salariales de entre 6% y 9% en esas regiones.

Hanson (2005) demuestra que los emigrantes tienen una selección positiva en cuanto a sus habilidades observables. Lo que indica que los trabajadores con mayor nivel de ingresos han contribuido a los diferenciales salariales regionales.

Gallardo Del Ángel (2023) es su trabajo sobre distribución del ingreso salarial usando técnicas paramétricas y no-paramétricas, y como parte de su análisis, estima la oferta laboral por grupos de trabajadores usando regresión Lowess para los años 2010 y 2020, confirmando la oferta laboral quebrada hacia atrás. Su análisis confirma que la oferta laboral tiene una parte elástica y otra inelástica.

Más se ha escrito sobre la precariedad del empleo y la desigualdad en el mercado laboral Mexicano. Anguiano Orozco y Ortiz Magallón (2013) argumentan que la reforma laboral del 2012 en México no creó empleos estables, provocó una precarización en el trabajo, legalizando el outsourcing y los contratos de capacitación inicial, ayudando a limitar los salarios y consolidando las condiciones precarias. Orlandina de Oliveira (2006) investiga la precariedad laboral juvenil en México a principios del siglo XXI utilizando la Encuesta Nacional de Juventud 2000. A través de un índice de precariedad/calidad del empleo y modelos de regresión, donde el estudio revela que el 31.3 % de los jóvenes ocupados se encuentra en condiciones muy críticas de trabajo. El artículo de Rojas García y Salas Páez (2008) analiza la precarización del empleo en México entre los años de 1995 y el 2004. Utilizando datos de la encuesta nacional de empleo y del IMSS. Este análisis revela la fuera polarización que se tiene en el mercado donde aparentemente el empleo con contrato permanente creció, los contratos verbales, la concentración de empleos precarios y microempresas son rasgos estructurales que se mantienen en el tiempo.

Saúl Escobar Toledo (2014) busca analizar la evolución que se tiene dentro de los salarios mínimos en México, mostrando una caída desde la década de los 80. A partir de los datos que se obtuvieron del Banco de México e INEGI, se demuestra que el salario mínimo cayó un 63.6 % entre los años de 1982 y 1992 y a partir de ahí se ha mantenido estancado hasta los años recientes. Liliana Meza González (2005) estudia la desigualdad salarial regional que se tiene en México entre los años 1988 y 1999, utilizando

datos obtenidos de la ENEU, revelando una tendencia al aumento de la desigualdad a nivel nacional, mostrando que este fenómeno se concentró en ciudades del centro y sur del país, mientras que al norte se observan reducciones. Graciela Bensusán (2006) ofrece un análisis crítico de la reforma laboral en México, donde señala que los cambios realizados fueron superficiales y no resuelven los problemas de desigualdad ni de precariedad. Juan Carlos Moreno-Brid et al. (2014) también analiza el salario mínimo de México comparándolo con otros países de Latinoamérica. Donde concluye que pase a la alta productividad que se tiene, los salarios reales han caído y la pobreza laboral ha aumentado. Graciela Bensusán (2013) evaluar la reforma laboral de 2012, descubriendo que si contribuye a un desarrollo incluyente en México. Y con un diagnóstico de oferta de trabajo, concluye que el centrarse en la flexibilidad y la subconcentración, no resuelve los problemas de informalidad que se tiene.

Minian y Martínez Monroy (2018) estudian lo vulnerable que se ha vuelto el empleo en México ante la automatización, utilizando la metodología ofrecida por Frey y Osborne (2013), concluye que el 63% del empleo total y el 64.5% del empleo manufacturero corren un alto riesgo de poder ser automatizados. El artículo de Orlandina de Oliveira (2009) retoma el tema de precariedad laboral juvenil, utilizando datos de la ENE 2004, encontró que el 31.3% de los jóvenes ocupados trabaja en condiciones críticas y que los factores más influyentes son el sector, el tamaño que tiene la empresa y la región.

En su artículo, Ocampo Merlo (2022) reconstruye el proceso de la reforma laboral de 2019 en México y analiza sus potenciales efectos en las relaciones laborales. Argumenta que la reforma fue impulsada por presiones externas, principalmente de T-MEC y que representa una ruptura con el modelo corporativista. Los cambios más significativos incluyen la desaparición de las juntas de conciliación y arbitraje. Aun que se reconocen avances, también se advierte que las estructuras de control empresarial siguen presentes por lo que el impacto real de la reforma aún está en disputa.

Existe también una literatura importante respecto al avance en la participación laboral femenina y de jóvenes. Bhalotra y Fernández (2024) estudian los factores detrás del aumento de la participación laboral femenina en México entre los años de 1960 y 2015. Usando una descomposición estadística, encuentra que a el nivel educativo explica el 43.3% del aumento después de 1970. También influyeron la disminución de la fertilidad y los cambios en los empleos disponibles. Los autores señalan que en los años noventa hubo crecimiento no explicado solo por educación o trabajo, lo que puede estar relacionado con cambios culturales y sociales sobre el papel de las mujeres.

El artículo de Pérez Cruz y Pinto Pérez (2020) investiga los factores que afectan en la integración laboral que tiene los egresados universitarios mexicanos. A los estudiantes de Medicina de la Universidad de

Colima se les aplicó un análisis econométrico con datos del panel, encontrando que la titulación especialmente mediante el EGEL mejora las probabilidades de conseguir un empleo. También destaca que las habilidades personales como el trabajo en equipo, el manejo de tecnologías, las prácticas profesionales son factores clave para tener éxito laboral. Lo que sugiere que las universidades deben adaptar sus programas educativos para responder mejor a las demandas del mercado laboral.

En su investigación, María del Carmen Salgado Vega (2005), analiza la relación que se tiene entre la educación superior y el mercado de trabajo en México, a través de su estudio de caso con egresados de la licenciatura en economía de la UAEM. Encontró que solo el 55% trabaja en algo relacionado a la carrera que estudiaron y que la mayoría lo hace en el sector de servicios, sobre todo en la administración pública y la educación. También encontró que 75.6% consiguió trabajo a través de contactos personales o familiares, concluyendo que, aunque estudiar da una ventaja, el valor de los títulos universitarios se ha reducido.

Enrique Hernández Laos (2004) examina el mercado laboral para profesionistas en México durante la década de 1990. Donde se basa en censos de la población entre los años 1990 y 2000 concluyendo que la oferta de egresados de educación superior creció a un ritmo de 6.7% anual, superando el crecimiento económico anual de 3.5%. El análisis revela una saturación de las ocupaciones profesionales, lo que llevó a un excedente de 880,000 profesionistas, que terminaron por emplearse en ocupaciones de una menor cualificación.

Este trabajo es una estimación de la oferta de trabajo en México usando datos de la Encuesta Nacional de Ocupación y Empleo (ENOE) para el primer trimestre de 2020. El objetivo es explicar la oferta laboral quebrada hacia atrás que predice la teoría y las estadísticas oficiales. Está organizado de la siguiente forma: La sección 1 presenta la INTRODUCCIÓN que contiene una revisión de la literatura sobre la oferta y situación del mercado laboral en México; la sección 2 expone el modelo clásico de oferta laboral; la sección 3 presenta los principales métodos de efectos ingreso y sustitución; la sección 4 trata la selección y regresión Lowess y la sección 5 describe los resultados obtenidos.

1. El modelo clásico de oferta laboral

A diferencia de una utilidad que depende únicamente del consumo de bienes, el individuo elige entre una canasta de bienes y el consumo de tiempo de ocio. Se asume que esta función de utilidad es dos veces diferenciable y que es semi-cóncava, lo que presupone una utilidad decreciente. Asumimos que en la canasta de consumo todos los bienes son normales, es decir, su consumo se incrementa cuando el

ingreso o el presupuesto de los individuos aumenta. El tiempo de ocio es el tiempo que no se dedica al trabajo (oferta laboral) y se usa para todo tipo de actividades que no son remuneradas en el mercado laboral, como puede ser hobbies, recreación, aprendizaje, descanso, alimentación, etc. Suponemos una función de utilidad del tipo Cobb-Douglas:

$$u(c, l) = c^\alpha l^\beta$$

Esta función de utilidad expresa las preferencias individuales respecto a dos bienes, ocio l y una canasta de bienes de consumo c . Podemos hacer una transformación monótonica de la función de utilidad y convertirla en logaritmos:

$$U(c, l) = \alpha \ln(c) + \beta \ln(l)$$

donde U está en términos de logaritmos, los coeficientes α y β representan los porcentajes de consumo de ocio y bienes¹ o la valuación individual respecto a su consumo. Una función de utilidad de este tipo presupone retornos constantes a escala donde $\alpha + \beta = 1$, y se asume que esta función de utilidad es representativa para todos los individuos. La restricción presupuestal representa la dotación de recursos iniciales, presentada por:

$$pc = wL + A$$

donde el gasto en consumo es pc y el precio de esa canasta de bienes está representado por p que es igual al ingreso total $wL + A$, tanto laboral como no laboral. Otra restricción importante que tienen los individuos es el tiempo total disponible que se usa para ocio l o trabajo:

$$T = L + l$$

donde L es la oferta de trabajo en términos de horas dedicadas a trabajar y T es el total de horas disponibles al día, semana, mes o año que es igual para todos los individuos. La tasa de salario es w y se paga en el mismo periodo que L , mientras que A es el ingreso no salarial de las personas o el ingreso adicional de otros miembros de un hogar.

Si se considera que el precio de los bienes c de consumo es un numerario igual a la unidad $p = 1$, entonces la tasa de salario real simplemente se denota como w . Los individuos eligen de manera óptima entre la cantidad de bienes de consumo c y la cantidad de ocio l que resulta en maximizar la función de utilidad con respecto a la dotación de recursos iniciales y el tiempo disponible. Pero como T es una cantidad fija, el ocio puede ser expresado también en términos de la cantidad de trabajo $l = T - L$. De

esta forma obtenemos la oferta laboral óptima que representa la cantidad de tiempo dedicada al trabajo y con ello se determina al mismo tiempo la cantidad óptima de ocio. La oferta laboral óptima nos dice que a medida que se incrementa el salario, de la misma forma se incrementa la oferta laboral:

$$L^* = \frac{\alpha wT - \beta A}{(\alpha + \beta)w}$$

En este modelo simple estático, los individuos son tomadores de precios y la tasa de salario w esta positivamente relacionada con la oferta laboral, es decir, una tasa de salario alta en exceso induciría una oferta excesiva de trabajo. Esta relación es generalmente aceptada y describe que en el caso de sociedades con trabajos y trabajadores homogéneos y libre competencia donde a mayor tasa salarial mayor será la oferta de trabajo. Sin embargo, en sociedades modernas complejas con una gran heterogeneidad de trabajadores y tipos de trabajos, los determinantes de la oferta laboral se vuelven más difíciles de identificar. Hay factores institucionales y características individuales que hacen más difícil observar la motivación para cada decisión de trabajar.

MATERIALES Y MÉTODOS

Efectos ingreso y sustitución

Asumiendo que los precios de los bienes de consumo permanecen sin cambio, así como el ingreso no salarial, se aprecia que la tasa de salario determina la pendiente de la restricción presupuestal de las personas, lo cual influye en la decisión de participar en el mercado laboral, así como el número de horas de trabajo a ofertar.

Por otro lado, se observa que el ingreso total posible de un individuo es igual a su gasto total $wT + A = wl + c$, es decir, el ingreso laboral que un individuo puede obtener por trabajar el total de horas disponibles y su dotación de ingreso no laboral tiene que ser igual a su consumo de bienes y su gasto en ocio. Es decir, tanto una hora de trabajo se valora a una tasa w que es la misma que se usa para valorar una hora de ocio. Se asume que la canasta de bienes de consumo como el ocio son bienes normales, por lo tanto, se espera que un incremento en la tasa de salario incremente el consumo de ocio. Esto se conoce como el efecto ingreso debido a un cambio en los precios relativos, en este caso un incremento en el presupuesto disponible.

Pero una tasa mayor de salario hace más caro el ocio o, en otras palabras, hace una hora de trabajo más remunerativa y los individuos encontrarán que es posible ahora trabajar más e incrementar su ingreso

TRANSICIÓN DE ESTADÍSTICA A CIENCIA DE DATOS

para consumir más bienes a costa de sacrificar tiempo de ocio. Este efecto de un cambio relativo en los precios que ocasione un incremento en el presupuesto y que permite sustituir el consumo de un bien por otro se conoce en la literatura económica como el efecto sustitución. Tanto el efecto ingreso como el de sustitución trabajan en direcciones opuestas cuando se trata del bien ocio. A pesar de que el ocio es un bien normal, la diferencia radica en que el individuo es productor (tiene una dotación de tiempo disponible) y consumidor al mismo tiempo. Para entender cómo los efectos sustitución e ingreso afectan de manera diferenciada a la oferta laboral podemos graficar estos efectos para dos individuos con preferencias diferentes:

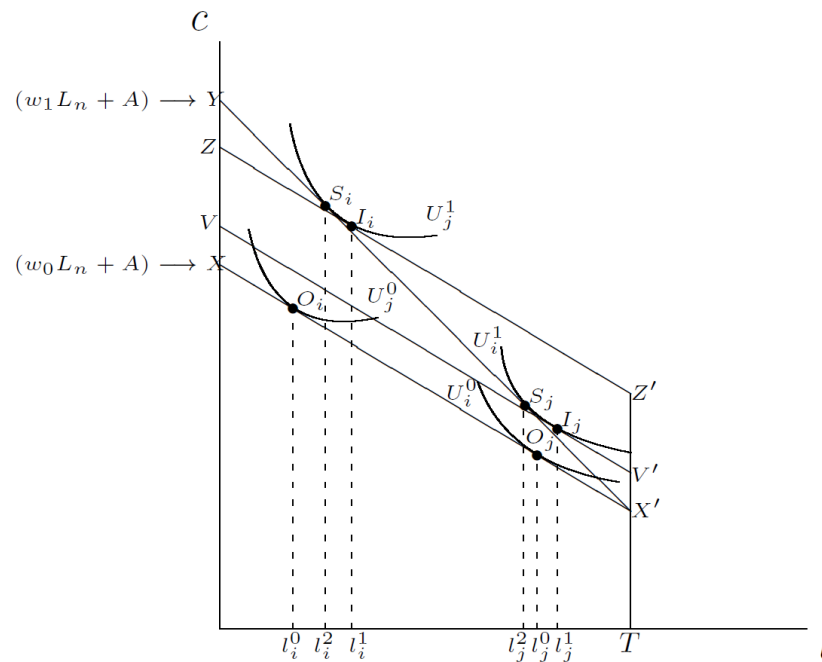


Gráfico 1. Efectos sustitución e ingreso.

El gráfico 1 presenta a dos individuos $n = i, j$ con las mismas restricciones presupuestales, pero con diferentes utilidades U_i y U_j . Ambos individuos tienen preferencias diferentes en cuanto a su consumo de bienes y su consumo de ocio. El individuo consume una cantidad menor de bienes, pero consume más ocio respecto al otro individuo, mientras que el individuo i prefiere consumir más bienes, eso significa sacrificar su tiempo de ocio. Ambos se encuentran inicialmente en la línea presupuestal XX' el primero en el punto óptimo O_j y el otro en el punto O_i . Ambos disfrutan de la tasa de salario w_0 y consumen de manera diferente dada sus preferencias individuales. Supongamos que por factores externos la tasa de salario se incrementa a w_1 , lo que coloca a ambos individuos en una línea

presupuestal más alta YX' . Para medir el efecto ingreso podemos trazar dos líneas paralelas a la línea de presupuesto original XX' para cada individuo, y estas líneas son VV' para j y ZZ' para i . Los puntos I_j e I_i representan el efecto ingreso: el cambio en la utilidad individual derivado del aumento salarial, manteniendo los precios relativos o el salario constante. El individuo i incrementa su consumo de ocio de l_i^0 a l_i^1 mientras que el individuo j lo incrementa de l_j^0 a l_j^1 .

El efecto sustitución se puede medir si permitimos el cambio en los precios relativos como en los puntos S_j y S_i en la línea YX' . Los individuos mantienen su nivel de utilidad después del cambio en el salario, pero ahora cada individuo reduce su consumo de ocio para consumir bienes en magnitudes diferentes. Por un lado, el individuo i disminuye su consumo de ocio de l_i^1 a l_i^2 , lo que le permite consumir aún más bienes al tener un cambio positivo en oferta laboral original. Por el contrario, el individuo j disminuye su consumo de ocio de l_j^1 a l_j^2 pero su oferta neta laboral disminuye si aún continúa consumiendo muchos más bienes. En suma, el incremento salarial tiene efectos diferentes en la oferta laboral de cada individuo, mientras que el individuo i disminuye su oferta laboral de $T - l_i^0$ a $T - l_i^2$, el individuo j la aumenta de $T - l_j^0$ a $T - l_j^2$.

Es posible ver que el efecto sustitución supera al efecto ingreso cuando se trata de personas con bajos salarios. Esto quiere decir que las personas con poco ingreso, y con una alta propensión a consumir, tienen una marcada preferencia por el consumo de bienes, o al menos, una necesidad que los impulsa a ofertar más mano de obra. Este efecto se revierte cuando se trata de altos salarios o con personas que tienen una gran preferencia por el ocio, donde el efecto sustitución es dominado por el efecto ingreso. De esta forma podemos observar que la curva de oferta laboral tiene una pendiente positiva para salarios bajos y después se hace negativa cuando la tasa de salario se hace muy grande, tal como se presenta de una forma general en el gráfico 2.

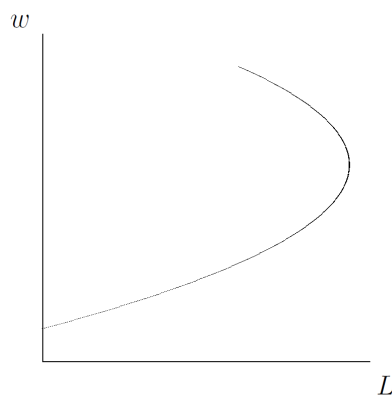


Gráfico 2. Oferta laboral hacia atrás.

La evidencia empírica también muestra que al menos para el grupo de trabajadores con más altos ingresos, la oferta laboral disminuye progresivamente. La tabla 1 muestra ofertas laborales promedio por grupo de ingresos para México. Se observa claramente que la oferta se incrementa cuando el ingreso laboral es mayor para los tres primeros grupos de ingresos hasta llegar a los diez pesos por hora. Al pasar a grupos de ingreso mayores, las horas trabajadas en promedio disminuyen, es decir, los trabajadores con salarios más altos tienden a trabajar menos que los trabajadores con salarios más bajos. Es posible que las preferencias individuales se reflejen en la curva de oferta laboral, en algún punto se hace inelástica y después se dobla hacia atrás. Por lo tanto, si algunos trabajadores presentan aumento en el salario no implica necesariamente un aumento en su oferta laboral.

Tabla 1. Promedio de horas trabajadas por rango de ingreso para México 2010.

Salario (pesos por hora)	Horas trabajadas (promedio por semana)
menos de \$1	40.5
\$1 - \$4.999	46.2
\$5 - \$9.999	49.9
\$10 - \$14.999	47.4
\$15 - \$19.999	45.9
\$20 - \$24.999	44.3
\$25 - \$29.999	41.1
Mayor de \$30	35.9

Fuente: Elaboración propia con datos de la ENOE 2010-Q1. INEGI.

Cabe recordar que el objetivo de este trabajo es estimar y validar esta oferta laboral quebrada usando técnicas paramétricas y no-paramétricas, y comparar las elasticidades de la oferta laboral considerando grupos de trabajadores de bajos o altos salarios, que se podrán distinguir en el punto donde la oferta laboral se quiebra o revierte en su pendiente.

Corrección por selección y regresión Lowess

El objetivo de este trabajo es encontrar la oferta laboral usando una regresión Lowess (locally weighted scatterplot smoothing) y poder estudiar el efecto ingreso y la curva de oferta hacia atrás que predice la teoría y la evidencia empírica. Decidimos tomar como datos de análisis la Encuesta Nacional de Ocupación y Empleo (ENOE) para el primer trimestre de 2020, justo al comienzo de la pandemia

COVID19, un periodo que al menos no sufrió de las perturbaciones posteriores en el mercado laboral. Pero antes de proceder debemos hacer una pausa y reconocer que existen datos perdidos para salarios en muchas observaciones, causados por problemas de selección. Para resolver esto procedimos a realizar una corrección de Heckman (o modelo de selección de Heckman) que es un método econométrico diseñado para corregir el sesgo de selección o el problema de muestras no aleatorias. En una regresión tradicional supone la estimación de una ecuación de salarios:

$$y_i = X_i\beta + \varepsilon_i$$

donde y_i representa el salario del individuo i , X_i es un vector de características observables (educación, experiencia, etc.) y ε_i es el término de error. Hay que resaltar que solo observamos y_i para aquellas personas que están trabajando o que participan en el mercado laboral. Entonces la participación laboral se puede representar mediante una variable latente y_i^* definida como:

$$y_i^* = Z_i\gamma + u_i$$

donde Z_i son variables que determinan la participación, tales como: la edad, número de hijos, educación del cónyuge, etc., y u_i es un término de error. El individuo trabaja si $y_i^* > 0$, es decir:

$$D_i = \begin{cases} 1 & \text{si } y_i^* > 0, & \text{trabaja} \\ 0 & \text{si } y_i^* \leq 0, & \text{no trabaja} \end{cases}$$

Por tanto, el salario y_i solo se observa cuando $D_i = 1$. Si los errores ε_i y u_i están correlacionados, entonces la muestra observada no es aleatoria, y los estimadores β de Mínimos Cuadrados Ordinarios (MCO) serán sesgados e inconsistentes. En 1979 James Heckman propuso una solución que recurre a una estimación en dos etapas para corregir este sesgo. El procedimiento es: primero estimar la probabilidad de participar en el mercado mediante un modelo Probit,

$$P(D_i = 1) = P(y^* > 0) = P(Z_i\gamma + u_i > 0) = \Phi(Z_i\hat{\gamma})$$

donde $\Phi(\cdot)$ es la función de distribución acumulada normal estándar. A partir de esta estimación, se calcula el inverso de la razón de Mills:

$$\lambda_i = \frac{\phi(Z_i\hat{\gamma})}{\Phi(Z_i\hat{\gamma})}$$

donde $\phi(\cdot)$ es la densidad normal estándar y $\Phi(\cdot)$ es su función de distribución normal acumulada.

En una segunda etapa se estima el salario esperado, condicionado a ser observado: $E[y_i|X_i, D_i = 1] = X_i\beta + \rho\sigma_\varepsilon\lambda_i$, donde $\rho = \text{corr}(\varepsilon_i, u_i)$ es la correlación entre los errores de ambas ecuaciones, y σ_ε es la desviación estándar del error en la ecuación de salarios. Por tanto, la ecuación a estimar es:

$$y_i = X_i\beta + \delta\lambda_i + \varepsilon_i$$

donde $\delta = \rho\sigma_\varepsilon$. Esta ecuación puede estimarse por MCO, añadiendo λ_i como variable adicional. De esta forma, usando una corrección de Heckman, se obtienen los estimadores consistentes y luego se imputaron los valores predichos para salarios en la muestra.

Una vez que se obtuvo una muestra corregida, se procede a realizar una no paramétrica Lowess (locally weighted scatterplot smoothing). La regresión Lowess, es un método de regresión no-paramétrica usado para modelar relaciones entre variables sin asumir una forma funcional específica y es muy útil para visualizar relaciones no lineales en los datos. Este método ajusta un modelo alrededor de cada punto usando las observaciones cercanas, dando una mayor ponderación a las que se encuentran más cerca y que, por lo tanto, tienen una mayor influencia en el ajuste. La regresión Lowess combina tres cosas: 1. una regresión polinomial, ya sea lineal o cuadrática; 2. una estimación de Mínimos Cuadrados Ponderados, y 3. un parámetro para suavizar (smoothing) que determina el tamaño del vecindario o puntos cercanos. En lugar de ajustar una sola función, el método Lowess ajusta muchas pequeñas regresiones alrededor del cada punto x_i . Para cada valor objetivo x_0 , queremos un estimador suavizado $\hat{y}_0$.

Para seleccionar el conjunto de observaciones cercanas (neighborhood) elegimos un subgrupo de datos alrededor de x_0 . El tamaño de este grupo está determinado por el parámetro de extensión span parameter $\alpha \in [0,1]$, que representa la proporción de los datos usados en cada regresión local. Luego se asignan pesos basados en la distancia, por ejemplo, para los puntos cercanos a x_0 se calculan ponderaciones usando el kernel tricúbico:

$$\omega_i(x_0) = \left(1 - \left|\frac{x_i - x_0}{d(x_0)}\right|^3\right)^3$$

donde $d(x_0)$ es la distancia desde x_0 hasta el punto más lejano en el conjunto de observaciones. Después usando el conjunto de datos se corre una regresión lineal usando Mínimos Cuadrados Ponderados:

$$\min_{\beta_0, \beta_1, \dots, \beta_p} \sum_i \omega_i(x_0) (y_i - \beta_0 - \beta_1(x_i - x_0) - \dots)^2$$

La idea es predecir el mejor ajuste en el punto x_0 para un modelo lineal local. Este proceso se repite hasta obtener una curva suavizada que mejor se ajusta a los datos. Finalmente se realiza una suavización de los valores encontrados con base en un parámetro suavizante α cuyos valores son escogidos usando validación cruzada (cross-validation). El método LOWESS reduce la influencia de los outliers, lo cual solo haría que el error fuera muy grande en un método de Mínimos Cuadrados Ordinarios tradicional. Este método paramétrico es flexible y se ajusta a los datos de una mejor manera que otros métodos paramétricos que asumen ya un modelo dado, describiendo las relaciones no lineales de una manera más efectiva.

RESULTADOS

Los gráficos 3 y 4 muestran las regresiones Lowess para salarios en función de horas de trabajo, para los dos grupos de trabajadores hombres y mujeres. Cabe recordar que la regresión Lowess es un método no-paramétrico donde los datos determinan el modelo, así que no estamos restringidos a estimar un modelo paramétrico determinado guiado por una distribución determinada. Las líneas de oferta laboral obtenidas siguen un patrón predicho por la teoría como en el gráfico 1, donde se muestra la oferta doblada hacia atrás: Existe un parte creciente y elástica respecto al salario, y otra parte decreciente e inelástica. Es decir, en la parte de salarios bajos, las horas de trabajo se incrementan cuando aumenta el salario, entonces la oferta laboral se quiebra y las horas de trabajo disminuyen cuando los salarios se hacen muy altos.

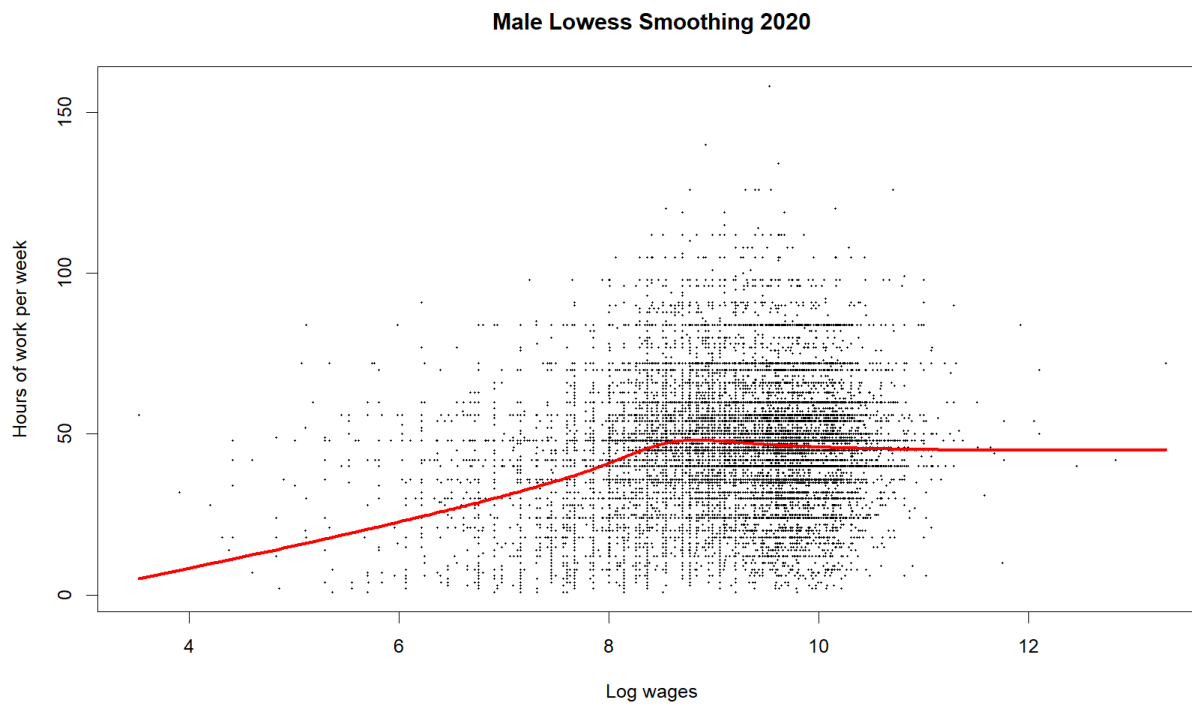


Gráfico 3. Regresión Lowess para trabajadores hombres 2020 Q-1

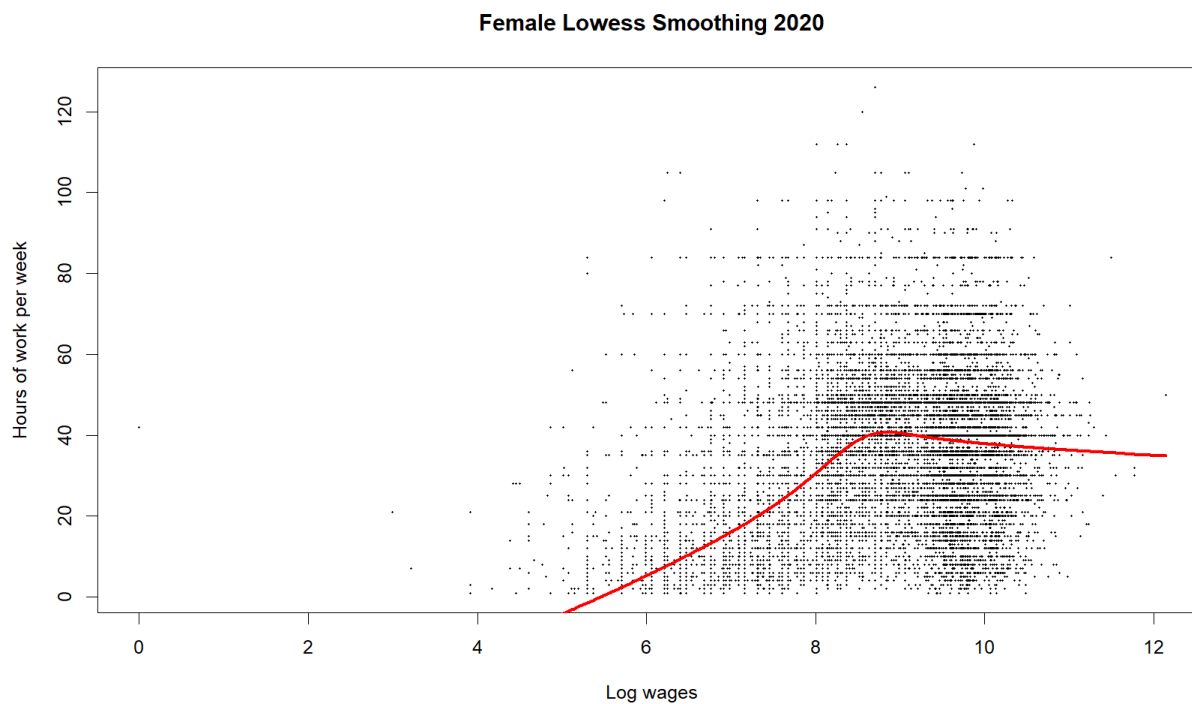


Gráfico 4. Regresión Lowess para trabajadores mujeres 2020 Q-1

Lo más relevante de la regresión Lowess es que permite localizar el punto de quiebre de la oferta laboral, es decir, el salario donde la curva de oferta laboral se dobla hacia atrás, y de esta forma podemos dividir la parte elástica con la inelástica de cada curva de oferta laboral, tanto para hombres como para mujeres. En el primer trimestre de 2020 el punto de quiebre para la oferta laboral femenina fue de 40.6 horas de trabajo a la semana, mientras que para los hombres el punto de quiebre fue de 48 horas a la semana. Esto no quiere decir que no hay mujeres que trabajaron más de 40.6 horas a la semana por un salario menor, o que existen hombres que trabajaron menos de 48 horas por un salario mayor. La regresión Lowess es un método de estimación que sigue la distribución de los datos, y en este caso, hace regresiones locales que se ajustan a la forma de los datos agrupados. Los outliers afectan la línea de regresión lineal debido a que contribuyen en demasía al error, pero en la regresión Lowess los puntos más alejados tienen un peso mucho menor, dejando que cada regresión local se ajuste mejor a los datos.

Una vez que se tienen los puntos de inflexión de cada oferta laboral para hombres y mujeres, se dividió cada muestra en dos, para separar trabajadores con bajos y altos salarios, con el objetivo de estimar las elasticidades para cada subgrupo. Después se realizó una regresión bietápica (Two Step Least Squares) para corregir por endogeneidad y se usaron variables instrumentales para estimar la variable salario. Finalmente se estimó la oferta laboral explicada por las variables: salario, nivel educativo y el ingreso de otros miembros del hogar. Las elasticidades estimadas aparecen en la tabla 2 al final del trabajo. Se puede observar que la evidencia empírica confirma lo predicho por la teoría clásica, donde los trabajadores hombres con los salarios bajos tienen una oferta elástica con un coeficiente de 1.46 mientras que los trabajadores con salarios altos tienen una oferta inelástica con un coeficiente de 0.13. Es decir, los trabajadores hombres con altos salarios tienen un incremento en su oferta laboral que es menos que proporcional al incremento al salario, pero sucede lo contrario con los trabajadores con bajos salarios. Para el caso de las mujeres, las elasticidades del salario tanto para bajos como altos ingresos inelásticas, aunque las mujeres con salarios más bajos siguen teniendo una oferta mucho más elástica que las trabajadoras con salarios más altos.

Para el caso de la variable educación, el cambio porcentual resultó pequeño y negativo para todos los grupos con excepción de los hombres con bajos salarios. Esto quiere decir que la mayoría de los trabajadores en México disminuyen su oferta laboral conforme más niveles de educación formal tienen, pues a mayor capital humano mayor será el salario y por lo tanto su consumo de ocio será mayor con una menor propensión a consumir bienes. Hay que destacar que en economías más desarrolladas donde

el sector servicios contribuye más al ingreso nacional, el concepto de ocio puede contener actividades que propician la acumulación de capital humano.

En cuanto a la variable ingreso del hogar, solo es estadísticamente significativa para los hombres, con una elasticidad negativa debido a que un aumento en otros ingresos por parte del conyugue, hace que disminuya su oferta laboral. Es interesante encontrar que la elasticidad del ingreso del conyugue no es estadísticamente significativa para las mujeres, lo que puede ser un indicio de mayor independencia económica.

Para concluir, podemos resaltar la importancia de diversos métodos no-paramétricos en el análisis económico. En este caso, se aplicó para la distinción de una curva de oferta que no es lineal y cuya estimación es difícil si usamos métodos tradicionales de regresión lineal.

REFERENCIAS

- Anguiano Orozco, A. Ortiz Magallón, R. 2013. Reforma laboral en México: precarización generalizada del trabajo Reforma laboral en México: precarización generalizada del trabajo. El Cotidiano 18295–104.
- Bensusán, G. 2006. La reforma laboral La reforma laboral. Economía UNAM 3933–53.
- Bensusán, G. 2013. Reforma laboral, desarrollo incluyente e igualdad en México Reforma laboral, desarrollo incluyente e igualdad en México. Naciones Unidas, CEPAL.
- Bhalotra, S. Fernández, M. 2024. El aumento de la participación laboral femenina en México: factores de oferta y demanda El aumento de la participación laboral femenina en México: factores de oferta y demanda. The World Bank Economic Review 382319–350.
- Blundell, R., Ham, J. Meghir, C. 1987. Unemployment and female labour supply Unemployment and female labour supply. The Economic Journal 44–64.
- Blundell, R. MaCurdy, T. 1999. Labor supply: A review of alternative approaches Labor supply: A review of alternative approaches. Handbook of labor economics 31559–1695.
- Bodie, Z., Merton, R.C. Samuelson, W.F. 1992. Labor supply flexibility and portfolio choice in a life cycle model Labor supply flexibility and portfolio choice in a life cycle model. Journal of economic dynamics and control 163427–449.
- Cahuc, P. 2004. Labor economics Labor economics. MIT press.
- Campos Vázquez, R. Esquivel, G. 2015. Informalidad, empleo y capital humano en México Informalidad, empleo y capital humano en México.

- Card, D. 1991. Intertemporal labor supply: An assessment Intertemporal labor supply: An assessment . National Bureau of Economic Research.
- Chiappori, PA. 1988. Rational household labor supply Rational household labor supply. *Econometrica: Journal of the Econometric Society*63–90.
- De Oliveira, O. 2006. Jóvenes y precariedad laboral en México Jóvenes y precariedad laboral en México. *Papeles de Población*124937–73.
- DeOliveira, O. 2009. El trabajo juvenil en México a principios del siglo XXI El trabajo juvenil en México a principios del siglo xxi.
- Escobar Toledo, S. 2014. Salarios mínimos: desigualdad y desarrollo Salarios mínimos: desigualdad y desarrollo. *Economía UNAM*113394–109.
- Gallardo Del Angel, R. 2023. Wage income distribution in Mexico: A nonparametric approach Wage income distribution in Mexico: A nonparametric approach. *Economía: teoría y práctica*5871–91.
- Gong, X. Van Soest, A. 2002. Estructura familiar y oferta laboral femenina en la Ciudad de México Estructura familiar y oferta laboral femenina en la ciudad de México. *Revista de Recursos Humanos*371163–191.
- González, LM. 2005. Mercados laborales locales y desigualdad salarial en México Mercados laborales locales y desigualdad salarial en México. *El Trimestre Económico*72285133–178.
- González, SA. 2017. La pobreza laboral en México La pobreza laboral en México. *Análisis Económico*328177–101.
- Ham, JC. 1982. Estimation of a labour supply model with censoring due to unemployment and underemployment. *The Review of Economic Studies*493335–354.
- Hanson, GH. 2005. Emigration, labor supply, and earnings in Mexico Emigration, labor supply, and earnings in Mexico.
- Heckman, J. 1974. Shadow prices, market wages, and labor supply Shadow prices, market wages, and labor supply. *Econometrica: journal of the econometric society*679–694.
- Heckman, JJ. 1979. Sample selection bias as a specification error Sample selection bias as a specification error. *Econometrica: Journal of the econometric society*153–161.
- Heckman, JJ. Sedlacek, GL. 1990. Self-selection and the distribu-tion of hourly wages Self-selection and the distribution of hourly wages. *Journal of Labor Economics*S329–S363.

- Hernández Laos, E. 2004. Panorama del mercado laboral de profesionistas en México Panorama del mercado laboral de profesionistas en México. Economía UNAM1298–109.
- Killingsworth, M. 1983. Labor Supply Labor supply. New YorkCam-bridge University Press.
- Killingsworth, MR. Heckman, JJ. 1986. Female labor supply: A survey Female labor supply: A survey. Handbook of labor economics1103– 204.
- Martínez, G. 2012. Estimación de las elasticidades de la oferta laboral en México Estimación de las elasticidades de la oferta laboral en México. Disponible en SSRN 2109933.
- Mincer, J. 1962. Labor force participation of married women: A study of labor supply Labor force participation of married women: A study of labor supply. Aspects of labor economics Aspects of labor economics (63–106). Princeton University Press.
- Mincer, J. 1970. The distribution of labor incomes: a survey with special reference to the human capital approach The distribution of labor incomes: a survey with special reference to the human capital approach. Journal of Economic Literature811–26.
- Mincer, J. 1994. The production of human capital and the lifecycle of earnings: Variations on a theme The production of human capital and the lifecycle of earnings: Variations on a theme . National Bureau of Economic Research.
- Mincer, J. 1996. Economic development, growth of human capital, and the dynamics of the wage structure Economic development, growth of human capital, and the dynamics of the wage structure. Journal of Economic Growth1129–48.
- Mincer, J. Polachek, S. 1974. Family investments in human capital: Earnings of women Family investments in human capital: Earnings of women. Marriage, family, human capital, and fertility Marriage, family, human capital, and fertility (76– 110). NBER.
- Minian, I. Martínez Monroy, A. 2018. El impacto de las nuevas tecnologías en el empleo en México El impacto de las nuevas tecnologías en el empleo en México. Problemas del desarrollo4919527–53.
- Moreno-Brid, JC., Garry, S. Monroy-Gómez-Franco, LA. 2014. El salario mínimo en México El salario mínimo en México. Economía UNAM113378–93.
- Ocampo Merlo, RE. 2022. Reforma laboral en México: una reconstrucción de su negociación y sus potencialidades prácticas Reforma laboral en México: una reconstrucción de su negociación y sus potencialidades prácticas. Espiral (Guadalajara) 29849–42.

- Pérez Cruz, OA. Pinto Pérez, R. 2020. Determinantes de la inserción laboral en egresados universitarios en México Determinantes de la inserción laboral en egresados universitarios en México. RIDE. Revista Iberoamericana para la Investigación y el Desarrollo Educativo 1121.
- Rojas García, G. Salas Páez, C. 2008. La precarización del empleo en México, 1995-2004 La precarización del empleo en México, 1995-2004. RELET-Revista Latinoamericana de Estudios del Trabajo 131939–78.
- Salgado Vega, MdC. 2005. Empleo y transición profesional en México Empleo y transición profesional en México. Papeles de Población 1144255–285.

Tabla 2: 2SLS Regresión por nivel de salarios 2020-Q1

Variables independientes	Hombres bajos salarios	Hombres altos salarios	Mujeres bajos salarios	Mujeres altos salarios
ln salario	1.4650 *** (0.1611)**	0.1348 *** (0.0482)**	0.6870 *** (0.2180)**	0.1167 * (0.0606)**
Nivel educativo	–0.0006 (0.0020)	–0.0062 *** (0.0014)**	–0.0085 ** (0.0040)**	–0.0096 *** (0.0014)**
ln ingreso del hogar	–0.7749 *** (0.1085)**	–0.0166 *** (0.0059)**	–0.1516 (0.1176)	–0.0066 (0.0057)
Constante	–1.8104 *** (0.3994)**	3.0351 *** (0.4058)**	–1.0907 (0.6646)	2.9484 *** (0.6050)**
Observaciones	6,891	5,270	9,552	2,966
R ²	–0.4847	–0.0738	0.3071	–0.0416
R ² ajustado	–0.4853	–0.0744	0.3069	–0.0427
Error estándar residual	0.6469 (df = 6887)	0.1842 (df = 5266)	0.6044 (df = 9548)	0.1751 (df = 2962)

Nota: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$ *

ANÁLISIS DE TENDENCIAS DE LA CONCENTRACIÓN DEL OXÍGENO DISUELTO Y DE LA TEMPERATURA DE LA REGIÓN HIDROLÓGICA GOLFO CENTRO, MÉXICO

Amy Yamileteh Loeza Beureth¹; María Teresa Leal Ascencio¹; Iván Ruiz Hernández¹;
Yesid Carvajal Escobar²

¹Universidad Veracruzana, México, ²Universidad del Valle, Colombia

RESUMEN

El objetivo del trabajo fue determinar si existen tendencias en la temperatura y en la concentración del oxígeno disuelto en los cuerpos de agua de la Región Hidrológico-Administrativa Golfo Centro, por lo que se analizaron los datos de calidad de agua de las 21 estaciones más completas que conforman la región, la cual abarca prácticamente todo el estado de Veracruz y partes de Oaxaca, Puebla e Hidalgo, resultando en que solamente 11 presentan tendencias estadísticamente significativas en la concentración del oxígeno disuelto y 18 en la temperatura del agua. Es de suma importancia el análisis de las series de datos de calidad de agua pues permite identificar problemas provocados por el Cambio Climático, así como planear y tomar las medidas necesarias para la adaptación de los sistemas acuáticos.

INTRODUCCIÓN

El término cambio climático denota un cambio en el estado del clima identificable a raíz de un cambio en el valor medio y/o en la variabilidad de sus propiedades y que persiste durante un periodo prolongado, generalmente cifrado en décadas o en periodos más largos. Denota todo cambio del clima a lo largo del tiempo, tanto si es debido a la variabilidad natural como si es consecuencia de la actividad humana (IPCC, 2014).

Conde (2006), menciona que el Sol es la fuente más importante para que funcione el sistema climático terrestre, por lo que, las variaciones de intensidad de la energía recibida, la distancia entre el sol y la tierra debido al movimiento de traslación y los cambios de órbita, la inclinación del planeta con respecto a su eje y su movimiento de rotación, hacen que no todo el año ni todos los años se reciba la misma cantidad de energía, lo que ha producido y producirá cambios climáticos naturales. Sin embargo, la presencia humana en este sistema es clara y evidente, ya que las emisiones antropogénicas recientes de gases de

efecto invernadero (GEI), como dióxido de carbono (CO_2), metano (CH_4) y óxido de nitrógeno (N_2O) son las más altas de la historia, pues han ido en aumento desde la era preindustrial, en gran medida como resultado del crecimiento económico y demográfico, pues los procesos industriales se han desarrollado básicamente quemando combustibles fósiles (petróleo, gas y sus derivados, como la gasolina), por lo que es sumamente probable que los efectos de estas emisiones y otros factores humanos como la deforestación generalizada, hayan sido la causa dominante del calentamiento observado a partir de la segunda mitad del siglo XX, pues cada decenio (1983-2012) ha sido sucesivamente más cálido en la superficie de la Tierra y es probable que este periodo haya sido el más cálido de los últimos 1,400 años en el hemisferio norte (IPCC, 2014), pues la temperatura global ha aumentado 0.74°C en los últimos 100 años y la tendencia de los últimos 50 años es de 0.13°C por década.

El IPCC (2007) señala que muchos sistemas hidrológicos están siendo afectados, pues hay un grado de confianza alto en que está aumentando la escorrentía y se está experimentando un adelanto de las fechas de caudal máximo primaveral en numerosos ríos alimentados por glaciares y por nieve, así como el calentamiento de lagos y de ríos en muchas regiones, lo que afecta su estructura térmica y la calidad del agua que contiene, ya que el aumento de temperatura acelera la evaporación, además, García *et al.* (2012) menciona que los impactos del cambio climático sobre los sistemas de agua dulce se deben a los incrementos observados y proyectados en la temperatura, nivel del mar y variabilidad de precipitación, lo que afecta la función y operación de la infraestructura hídrica existente, así como las prácticas de gestión integral de los recursos hídricos.

Asimismo, como resultado del aumento de la temperatura ambiental se desprenden diversas consecuencias para la calidad del ambiente acuático que desencadenan reacciones muy variadas como la aceleración de reciclamiento y disponibilidad de nutrientes (eutrofización), la reducción del oxígeno disuelto en el agua, mayor evaporación, incremento de la toxicidad de algunos metales (Rathore y Khangarot, 2002), modificación del equilibrio de especies acuáticas, aumento en la solubilidad de compuestos orgánicos, y mayores retos y costos en los tratamientos como la potabilización (Leal, 2012).

La temperatura del agua corriente es uno de los parámetros más importantes en los estudios de ecosistemas acuáticos, pues influye en los procesos químicos presentes en los sistemas fluviales, así como en las condiciones y el comportamiento biológico (Bates *et al.*, 2008). Park *et al.* (2010) señala, además, que estos cambios pueden influir directamente en los parámetros de calidad que dependen de ésta, incluidos el oxígeno disuelto, los potenciales de oxidorreducción, el pH, la estratificación y la actividad microbiana. De igual forma, el oxígeno disuelto es un indicador por excelencia, el cual define la calidad del agua para

la vida acuática, ya que es esencial para todas las formas de vida, especialmente para los organismos encargados de los procesos de autopurificación en aguas naturales (Kannel *et al.*, 2007), varía con la temperatura, salinidad, turbulencia, actividad fotosintética y presión atmosférica (Jackson *et al.*, 2001).

En México se han realizado muy pocos estudios sobre tendencias de temperaturas y concentración de oxígeno disuelto en cuerpos de agua, a pesar de que su análisis es de suma importancia, ya que permite identificar problemas provocados por cambio climático, así como planear y tomar las medidas necesarias para la adaptación de los sistemas acuáticos.

METODOLOGÍA

La Región Hidrológico-Administrativa (RHA) X Golfo Centro, ubicada al sureste de México, abarca una superficie de 102,354 km² en 432 municipios de los estados de Veracruz, Hidalgo, Puebla y Oaxaca (Conagua, 2018). Está conformada por las regiones hidrológicas (RH) 27, 28 y 29, Norte de Veracruz o Tuxpan-Nautla, Papaloapan y Coatzacoalcos, respectivamente, las cuales incluyen 11 de los 14 ríos principales que drenan el estado de Veracruz y desembocan en el Golfo de México (figura 1).

El objetivo del análisis estadístico es encontrar evidencia de una posible tendencia significativa que indique que la RHA Golfo Centro presenta un posible cambio climático. Tomando en cuenta lo anterior, se recopilieron los datos de las estaciones de calidad de agua de las regiones hidrológicas 27, 28 y 29, los cuales fueron proporcionados por la Dirección Técnica del Organismo de Cuenca X: Golfo Centro, de la Comisión Nacional del Agua (CONAGUA). Se extrajeron los datos de la concentración de oxígeno disuelto en mg/L y de la temperatura del agua en °C por estación y posteriormente, se organizó la base por cuenca, estación y fecha. Una vez organizada la matriz, se evaluó la factibilidad de los datos de las estaciones mediante una valoración de los mismos para que de esta manera se tuvieran estaciones con un registro con al menos 90 datos, tomando en cuenta la fecha de inicio y el último muestreo, así como la cantidad de datos faltantes. Las estaciones seleccionadas se muestran en la tabla 1.

Figura 1. RHA X Golfo centro, ríos principales y estaciones de estudio. Fuente: Elaboración propia.

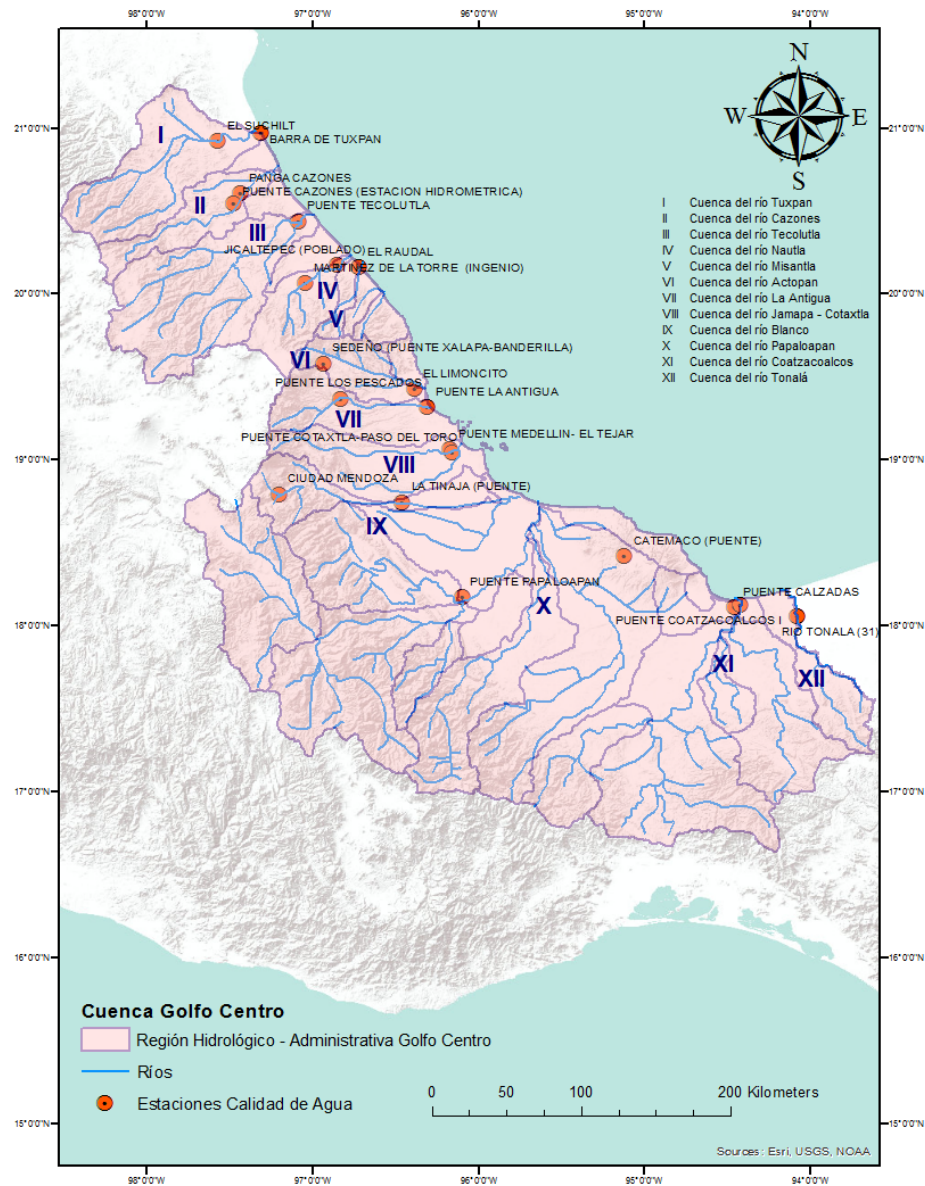


Figura 2. RHA X Golfo centro, ríos principales y estaciones de estudio.

Fuente: Elaboración propia.

Tabla 1. Estaciones seleccionadas por cuenca

Cuenca	Estaciones seleccionadas
Río Tuxpan	El Suchitl Puente Tuxpan
Río Cazones	Panga Cazones Puente Cazones
Río Tecolutla	Puente Tecolutla
Río Nautla	Jicaltepec Puente Martínez de la Torre
Río Misantla	El Raudal
Río Actopan	El Limoncito Puente Xalapa - Banderilla
Río La Antigua	Puente La Antigua Puente Los Pescados
Río Jamapa - Cotaxtla	Puente Medellín – El Tejar Puente Cotaxtla
Río Blanco	Puente La Tinaja Ciudad Mendoza
Río Papaloapan	Puente Papaloapan Laguna de Catemaco
Río Coatzacoalcos	Puente Calzadas Puente Coatzacoalcos I
Río Tonalá	Puente Tonalá

Para determinar la existencia de tendencias en las series de datos de temperatura y oxígeno disuelto, se realizaron análisis descriptivos para observar gráficamente la distribución, los datos atípicos y la relación entre las variables, y posteriormente, se aplicó la prueba no paramétrica de Mann-Kendall, ya que, a diferencia de las pruebas paramétricas, es menos sensible a los valores atípicos pues analiza los rangos de los valores registrados y no se ve afectado por la distribución de estos. El valor de significancia que se tomó en cuenta fue $\alpha=0.05$, donde:

Ho = No hay tendencia

Ha = Existe tendencia (positiva o negativa)

RESULTADOS

Los resultados muestran que, en la mayoría de las estaciones, los datos atípicos de la concentración de oxígeno disuelto no logran afectar que sigan una distribución normal, sin embargo, en los diagramas de cajas y bigotes se aprecia que en algunas estaciones los datos tienen rangos mucho más amplios que en otras, pues en algunas, el rango que va desde los 0 mg/L hasta los 16 mg/L. Asimismo, el rango de la temperatura del agua observada en las diferentes estaciones es muy grande, pues los valores mínimos se

encuentran por debajo de los 5°C y los máximos, cercanos a 40°C. A manera de ejemplo se muestran las figuras 2 y 3 en las que se observan los rangos de las variables y la simetría de las cajas.

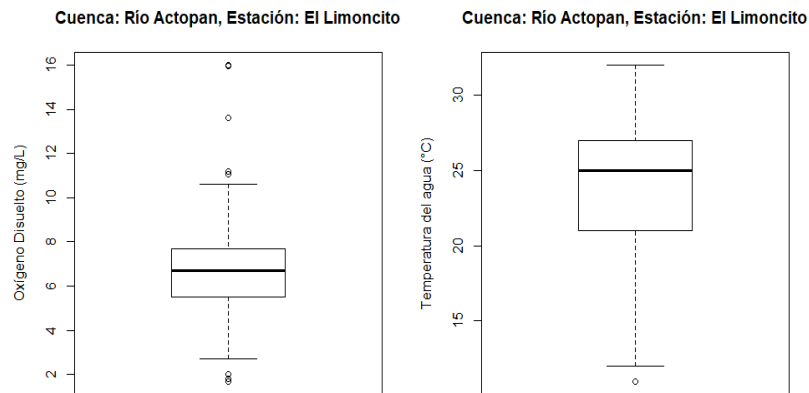


Figura 2. Diagramas de cajas y bigotes de la concentración del oxígeno disuelto y de la temperatura en la estación El Limoncito, cuenca río Actopan.

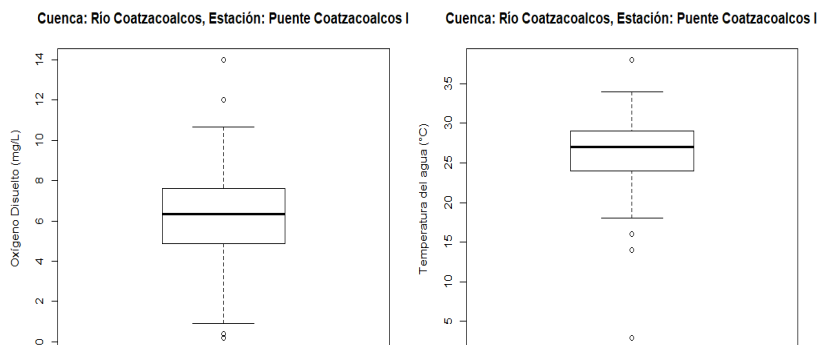


Figura 3. Diagramas de cajas y bigotes de la concentración del oxígeno disuelto y de la temperatura en la estación Puente Coatzacoalcos I, cuenca río Coatzacoalcos.

Asimismo, para conocer si existía correlación entre las variables, se calculó el coeficiente de correlación de Pearson, el cual fue muy cercano a 0 para todas las estaciones, lo que evidencia que el oxígeno disuelto no depende única y linealmente de la temperatura del agua. Sin embargo, en las gráficas se pudo observar que, en 12 de las 21 estaciones, la relación entre las variables fue negativa.

Los p-value de la prueba estadística Mann-Kendall (tabla 2), muestran que el oxígeno disuelto de las estaciones El Suchitl, Jicaltepec, Puentes Tecolutla, Xalapa-Banderilla, La Antigua y Los Pescados,

Medellín-El Tejar, Cotaxtla, La Tinaja, Cd. Mendoza y Papaloapan, presenta una tendencia significativa y que, excepto en la estación Jicaltepec de la cuenca del río Nautla, la tendencia es negativa, es decir, la concentración del oxígeno ha ido en decremento a través de los años, como se observa en las figuras 4 y 5. Por otra parte, para la temperatura del agua (tabla 3), la prueba señala que las únicas estaciones en las que no se observa una tendencia significativa son El Suchitl, Puente Calzadas (figura 7) y Puente Tonalá, de las cuencas de los ríos Tuxpan, Coatzacoalcos y Tonalá, respectivamente. En la figura 6 se observa un ejemplo de tendencia positiva de la temperatura en la estación El Raudal, de la cuenca del río Misantla.

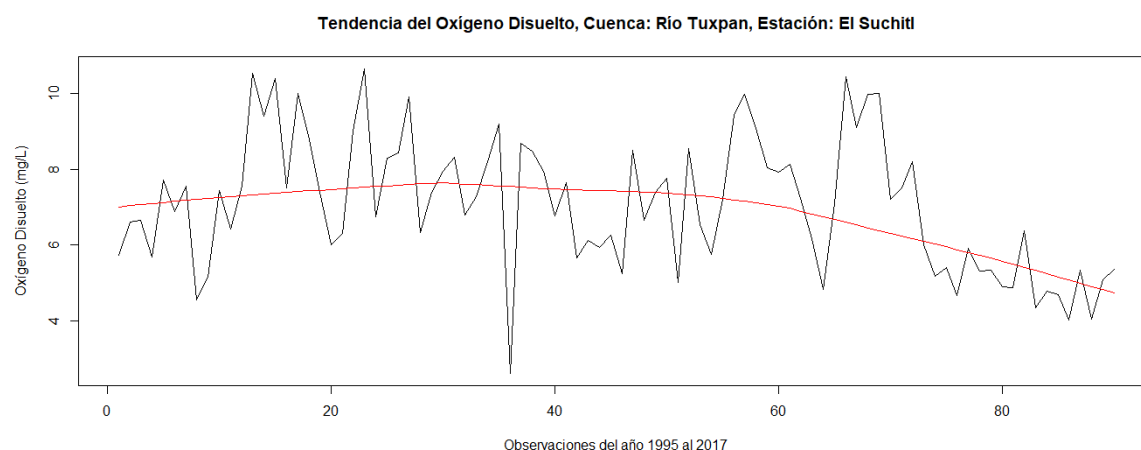


Figura 4. Ejemplo de tendencia negativa del oxígeno disuelto en la estación El Suchitl, cuenca río Tuxpan

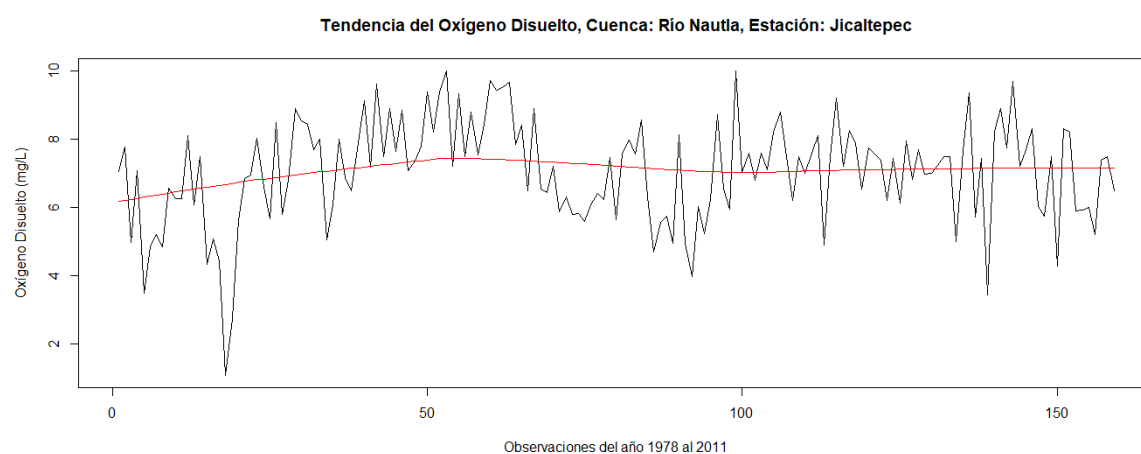


Figura 5. Ejemplo de tendencia positiva del oxígeno disuelto en la estación Jicaltepec, cuenca río Tuxpan

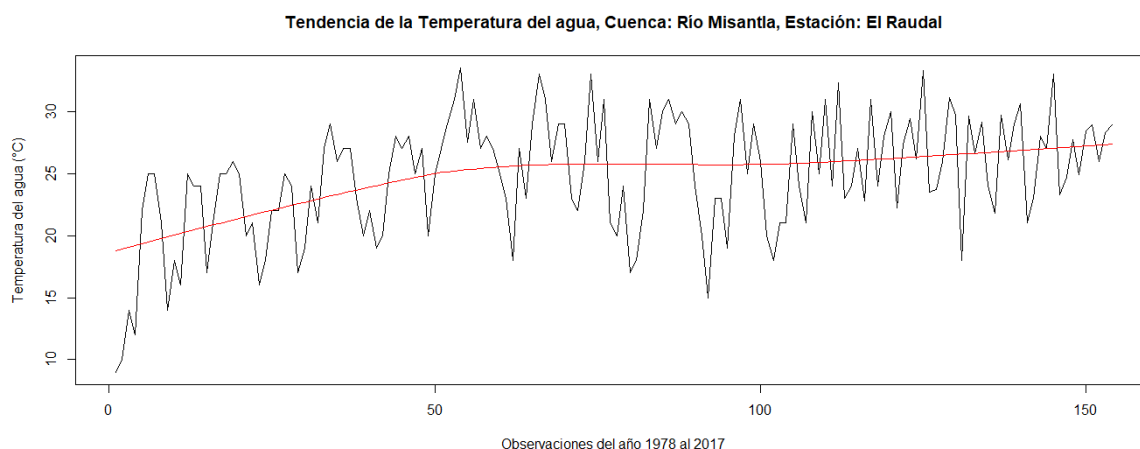


Figura 6. Ejemplo de tendencia positiva de la temperatura en la estación El Raudal, cuenca río Misantla

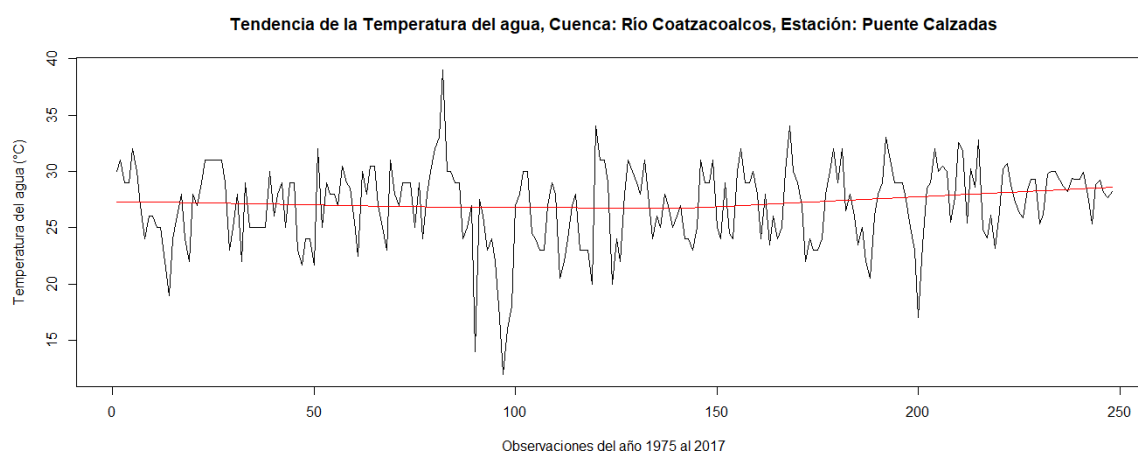


Figura 7. Ejemplo de datos sin tendencia significativa en la estación Puente Calzadas, cuenca río Coatzacoalcos

Tabla 2. Resultados de la prueba estadística Mann-Kendall del oxígeno disuelto por estación

Cuenca	Estaciones seleccionadas	Estadístico Mann-Kendall	p-value
Río Tuxpan	El Suchitl	-0.236	0.0010
	Puente Tuxpan	-0.082	0.0640
Río Cazones	Panga Cazones	-0.005	0.8989
	Puente Cazones	-0.017	0.7838
Río Tecolutla	Puente Tecolutla	-0.163	7.77×10^{-5}
Río Nautla	Jicaltepec	0.126	0.0073
	Puente Martínez de la Torre	-0.009	0.8669
Río Misantla	El Raudal	0.042	0.4333
Río Actopan	El Limoncito	-0.020	0.6074
	Puente Xalapa - Banderilla	-0.141	0.0083
Río La Antigua	Puente La Antigua	-0.102	0.0133

	Puente Los Pescados	-0.130	0.0023
Río Jamapa - Cotaxtla	Puente Medellín – El Tejar	-0.129	0.0021
	Puente Cotaxtla	-0.232	7.02×10^{-8}
Río Blanco	Puente La Tinaja	-0.126	0.0026
	Ciudad Mendoza	-0.271	1.25×10^{-10}
Río Papaloapan	Puente Papaloapan	-0.113	0.0102
	Laguna de Catemaco	0.009	0.8809
Río Coatzacoalcos	Puente Calzadas	-0.015	0.7299
	Puente Coatzacoalcos I	-0.041	0.3368
Río Tonalá	Puente Tonalá	-0.089	0.0843

Tabla 3. Resultados de la prueba estadística Mann-Kendall de la temperatura del agua por estación

Cuenca	Estaciones seleccionadas	Estadístico Mann-Kendall	p-value
Río Tuxpan	El Suchitl	0.069	0.3427
	Puente Tuxpan	0.200	1.19×10^{-5}
Río Cazones	Panga Cazones	0.197	3.46×10^{-6}
	Puente Cazones	0.252	7.87×10^{-5}
Río Tecolutla	Puente Tecolutla	0.270	$< 2.22 \times 10^{-16}$
Río Nautla	Jicaltepec	0.307	$< 2.22 \times 10^{-16}$
	Puente Martínez de la Torre	0.213	0.0002
Río Misantla	El Raudal	0.261	2.38×10^{-6}
Río Actopan	El Limoncito	0.331	$< 2.22 \times 10^{-16}$
	Puente Xalapa - Banderilla	0.125	0.0238
Río La Antigua	Puente La Antigua	0.303	$< 2.22 \times 10^{-16}$
	Puente Los Pescados	0.319	$< 2.22 \times 10^{-16}$
Río Jamapa - Cotaxtla	Puente Medellín – El Tejar	0.246	$< 2.22 \times 10^{-16}$
	Puente Cotaxtla	0.264	$< 2.22 \times 10^{-16}$
Río Blanco	Puente La Tinaja	0.170	9.50×10^{-5}
	Ciudad Mendoza	0.173	9.35×10^{-5}
Río Papaloapan	Puente Papaloapan	0.201	1.31×10^{-5}
	Laguna de Catemaco	0.226	0.0001
Río Coatzacoalcos	Puente Calzadas	0.074	0.0898
	Puente Coatzacoalcos I	0.121	0.0062
Río Tonalá	Puente Tonalá	0.047	0.3762

CONCLUSIONES

Es clara la tendencia de calentamiento de los cuerpos de agua de la Región Hidrológico-Administrativa Golfo Centro, pues 18 de las 21 estaciones mostraron una tendencia positiva. Por otro lado, solamente en 10 estaciones se muestra una tendencia negativa en la concentración del oxígeno disuelto y en una, positiva; lo cual confirma que la temperatura no es la única variable que influye en la concentración del oxígeno disuelto y tampoco lo hace de manera lineal, como indican los coeficientes de correlación de Pearson. Es necesario seguir el análisis para determinar la magnitud del cambio en las variables, es decir, en cuántos grados está aumentando la temperatura del agua y en cuántos mg/L está disminuyendo la concentración del oxígeno disuelto para conocer las consecuencias que estos cambios provocan tanto en la vida de los

sistemas acuáticos como en la población que usa y consume el agua estas cuencas hidrográficas, para poder planear y tomar las medidas necesarias de adaptación. Asimismo, se recomienda evaluar las tendencias de todas las cuencas del país para conocer la magnitud del problema.

REFERENCIAS

- Bates, B. C., Kundzewicz, Z. W., Wu, S., y Palutikof, J. P. (2008). *Climate Change and Water: Technical Paper of the Intergovernmental Panel on Climate Change*. Geneva: IPCC Secretariat.
- Comisión Nacional del Agua, 2018. Estadísticas del Agua en México. Recuperado de https://www.gob.mx/cms/uploads/attachment/file/259369/_2005_EAM2005.pdf, en marzo de 2019.
- Conde, C. (2006). México y el cambio climático global. Ciudad de México: Dirección General de Divulgación de la Ciencia de la Universidad Nacional Autónoma de México.
- García, M. C., Piñeros Botero, A., Bernal Quiroga, F. A., y Ardila Robles, E. (2012). Variabilidad climática, cambio climático y el recurso hídrico en Colombia. *Revista de Ingeniería*, 60-64.
- Intergovernmental Panel on Climate Change. (2007). *Cambio climático 2007: Informe de síntesis*. Ginebra, Suiza: Grupo Intergubernamental de Expertos sobre el Cambio Climático.
- Intergovernmental Panel on Climate Change. (2014). *Cambio climático 2014: Informe de síntesis*. Ginebra, Suiza: IPCC. Grupo Intergubernamental de Expertos sobre el Cambio Climático.
- Jackson, D. A., Peres-Neto, P. R., y Olden, J. D. (2001). What controls who is where in freshwater fish communities: the roles of biotic, abiotic and spatial factors. *Canadian Journal of Fisheries and Aquatic Sciences*, Vol. 58, 157-170.
- Kannel, P. R., Lee, S., Lee, Y., Kanel, S. R., y Khan, S. P. (2007). Application of Water Quality Indices and Dissolved Oxygen as Indicators for River Water Classification and Urban Impact Assessment. *Environmental Monitoring and Assessment*, 132, 93-110.
- Leal, M. T. (2012). Desarrollo de la guía metodológica para el análisis de la calidad del agua respecto al cambio climático. Xalapa, Veracruz.
- Park, J.-H., Duan, L., Kim, B., Mitchell, M. J., y Shibata, H. (2010). Potential effects of climate change and variability on watershed biogeochemical process and water quality in Northeast Asia. *Environment International*, 36, 212-225.
- Rathore, R. S., y Khangarot, B. (2002). Effects of Temperature on the Sensitivity of Sludge Worm *Tubifex tubifex* Müller to Selected Heavy Metals. *Ecotoxicology and Environmental Safety* 53, 27-36.

CORRELACIÓN ESPACIAL DE ACTIVIDADES ECONÓMICAS DE XALAPA BASADO EN EL ANÁLISIS CLÚSTER Y DE VALOR ATÍPICO (I ANSELIN LOCAL DE MORAN)

Juan Andrés Sánchez García, Angel Fernando Argüello Ortiz

Universidad Veracruzana, México

andressg89@hotmail.com, aarguello@uv.mx

RESUMEN

La ciudad Contemporánea representa un reto intelectual que desafía la inteligencia y la imaginación, su comprensión obliga a derribar las barreras disciplinarias y tender un puente entre los saberes por lo que estimula la creatividad para la búsqueda de métodos de observación, medición, representación y análisis. Un problema metodológico constante es la falta de modelos que determinen y representen los fenómenos transversales y longitudinales que transcurren en la ciudad, cuya visión única del urbanismo genera una limitante para evidenciar estos procesos, por lo que ésta condicionante disciplinar genera la necesidad de voltear a ver tendencias, conceptos y metodologías de otras disciplinas que gestionen el eje rector del urbanismo.

Este trabajo tiene como objetivo generar un análisis Geoestadístico que permita evidenciar la organización, correlación y asociación espacial atendiendo a conglomerados o grupos mediante el Análisis de clúster y de valor atípico, ya que este proceso identifica puntos calientes, puntos fríos y valores atípicos espaciales estadísticamente significativos mediante la estadística de I Anselin local de Morán; este proceso también permite gestionar e identificar zonas de concentración y dispersión que fomenten la base para localización de zonas vulnerables y soluciones estratégicas del desarrollo económico.

Palabras Clave: Correlación Espacial, Unidades Económicas, Índice de Morán

INTRODUCCIÓN

Las propuestas en investigación urbanística exigen una interdisciplinaria para explicar los fenómenos que acontecen dentro de la ciudad. El urbanismo es una disciplina que se encarga de estudiar los procesos de la ciudad, sin embargo, la nueva propuesta interdisciplinar otorga una visión contemporánea de enlazar disciplinas para explicar y representar de manera más adecuada los fenómenos urbanos. Con ello no es

necesario evidenciar las limitantes que el urbanismo ofrece sino explorar nuevas herramientas y metodologías que agrupan y articulan interpretaciones de la ciudad desde otras disciplinas.

Uno de los ejes del estudio de la ciudad es el relacionado con las configuraciones espaciales: concentraciones, dispersiones, zonas vulnerables, cohesiones, consolidaciones, estructuras, organizaciones etc., por lo que la grafía en urbanismo es una herramienta necesaria para interpretar los fenómenos urbanos y la manera en que se comporta el sistema de ciudad.

Por otra parte, una de las variables que ha modificado la fisionomía de la ciudad de Xalapa y que impulsa su crecimiento es el relacionado con la economía y que, aunque existe una visión desde la Economía Urbana, lo que el trabajo pretende es vincular al territorio con la economía desde un enfoque geoestadístico, el cual permite ser interpretado por el urbanismo.

Este trabajo tiene como objetivo mostrar, en primera instancia, la significancia de Unidades Económicas que tienen mayor peso para explicar un patrón de emplazamiento dentro de la ciudad, por lo que es necesario utilizar técnicas estadísticas para depurar la cantidad de observaciones Económicas y presentar un modelo de comportamiento espacial.

Por otra parte, ante los obstáculos que muestra el mapear concentraciones y dispersiones de Unidades Económicas ante una segmentación administrativa dada, es necesario romper el esquema morfológico de las AGEB y mostrar una disposición homogénea que represente mejor una fisionomía espacial.

Finalmente se pretende que esta nueva segmentación espacial contribuya al estudio de su morfología espacial y su distribución basado no solo en la concentración de Unidades Económicas, sino también en la correlación que existe entre celdas vecinas para obtener un patrón de emplazamiento mediante el Índice I de Morán. Dicho esto, se presenta una oportunidad de interpretar procesos geoestadísticos desde la perspectiva urbana que muestren el conocimiento interdisciplinar y de frontera que las investigaciones exigen.

Lo debatible es presentar distintas maneras de ver esta relación con diferentes técnicas para comparar interpretaciones, ya que en algunas se puede obviar y en otras se puede evidenciar tendencias. Este trabajo termina con una propuesta del I de Morán que pretende no solo ver la disposición espacial de las concentraciones económicas sino dialogar con correlaciones por cercanía de cada una de ellas, dejando entrever nuevas tendencias y patrones espaciales.

DESARROLLO

Estudiar la ciudad siempre ha sido un punto de discusión de investigadores inmersos en la disciplina del urbanismo y la arquitectura; la gama de posibilidades, realidades, escenarios, enfoques y fenómenos que presenta ésta conlleva un gran esfuerzo intelectual por entender la dinámica urbana.

La ciudad no es un objeto estático al cual se le debieran asignar problemas sino que existen diversos fenómenos que necesitan ser estudiados para comprender su funcionamiento, características y patrones, por lo que la multidisciplinariedad de los estudios puede ayudar a comprender una pequeña parte de la complejidad de la ciudad, ya que como refiere García Vázquez (2016), “la ciudad contemporánea es una criatura incierta, y hasta cierto punto vasto y complejo” y se considera la existencia de una dificultad de analizar un objeto tan grande, distinto y que crece a diferentes ritmos, atendiendo las circunstancias, situaciones y acontecimientos que interfieren en ella.

Resulta un problema metodológico la falta de modelos que determinen y representen los fenómenos transversales y longitudinales que transcurren en la ciudad, cuya visión única del urbanismo genera una limitante para evidenciar estos procesos, por lo que ésta condicionante disciplinar genera la necesidad de voltear a ver tendencias, conceptos y metodologías de otras disciplinas que gestionen el eje rector del urbanismo.

Un fenómeno recurrente en los estudios de la ciudad son los procesos de urbanización donde la ciudad de Xalapa no está exenta, por lo que, como algunas otras ciudades latinoamericanas, presenta situaciones y fenómenos que necesitan ser estudiados para entender sus procesos. Sería un error seguir designando el concepto de problema a los acontecimientos de la ciudad, por lo que se pretende denominarlos fenómenos ya que no es el interés de este trabajo resolverlos sino entenderlos.

Un fenómeno que se ha presentado en la ciudad de Xalapa desde mediados del siglo XX es la relación que se produce entre la población de las ciudades y la distribución de sus fuentes de trabajo, lo que ha desencadenado algunos análisis de movilidad, transporte, vivienda, segregación, cambios de uso de suelo, entre otros. No se debe olvidar que la ciudad está constituida de relaciones y de estructuras muy particulares, es decir de una organización mediante la cual funcionan muchos de los procesos dentro de ella, por lo que sería otro error generar una propuesta de realidad objetiva cuando es más adecuado mirar y evidenciar una realidad muy particular de la ciudad desde una perspectiva muy concreta.

La ciudad siempre está expuesta a cambios en los sistemas que predominan en las dinámicas espaciales. Uno de los más importantes es el que se refiere a las dinámicas económicas, puesto que plantean un vínculo muy cercano al desarrollo de las ciudades.

Uno de los supuestos más precisos que ayuda a entender la manera en que afecta la economía en el desarrollo en las ciudades es la Teoría Económica de Desarrollo, TDE en adelante, cuyo objetivo “es desentrañar las causas, los mecanismos y las consecuencias del crecimiento económico de los países” (Bustelo, 1998, pág. 19) y, sin ser economista ni pretender acercarse a la disciplina de la Economía, es necesario plantear la forma en que el sistema económico puede repercutir en el desarrollo de la ciudad.

Es prudente enmarcar una estrecha relación entre la Economía y el Territorio. Cuando se habla de ciudad se hace aparecer a la Economía Regional como “la Economía General aplicada a una determinada zona geográfica” (W. Richarson, 1973), es decir que una parte de la Economía se enfoca en el planteamiento espacio-tiempo ya que es necesario colocar estas variables para el análisis de los Sistemas Económicos.

La Economía Espacial, que aglutina la Economía Regional y la Economía Urbana, plantea una Teoría General de Localización, la cual tiene como uno de sus propósitos el “explicar el impacto que producen los cambios en las técnicas, en los costes de transporte, en los niveles de renta y gustos, etc., y sobre las estructuras de localización del consumo y de la producción” (íbid, pág. 112)

Aunque no es directamente el objetivo plantear una teoría de Economía Espacial, si es necesario referir al Economista Alemán August Lösch como un personaje que muestra aportes necesarios para este trabajo, por ejemplo: determinó al espacio como una variable fundamental para su estudio planteando el concepto de *región económica* que, según él, correspondía al resultado de una concentración espacial de las actividades con base en las redes que se forman de una producción común, diferenciando el supuesto de reparto homogéneo con los factores productivos y de las poblaciones a lo largo del espacio con el que comenzó a realizar sus trabajos; es esta homogeneidad parte de sus principales críticas al no ser tan general y ver su mayor representatividad en las actividades que son de servicio y no de actividades primarias.

El estructuralismo económico, a diferencia de la escuela keynesiana, neoliberalista, neoclásica o mercantilista, propone que las economías latinoamericanas necesitan ser renovadas al igual que el sistema capitalista internacional para lograr que los países pobres obtengan más ganancias del comercio entre países.

Claro está que las contribuciones en el sistema capitalista funcionan como un gran reloj donde cada pieza debe funcionar de la mejor manera para seguir en marcha, de tal forma que todos los actores y los sectores de la sociedad garanticen la competitividad de las relaciones comerciales y continúen los flujos financieros para el desarrollo de una ciudad.

Retomando las palabras de Jordi Borja (2002) en la búsqueda de la relación sistémica de la ciudad y su economía menciona que:

Si la ciudad es lo que se supone que debe ser, un lugar y una comunidad, con capacidad de autogobierno, hoy es también un territorio articulado que en las regiones más desarrolladas y urbanizadas convierte a la ciudad en un sistema de ciudades. Si este espacio tiene *vigencia económica*, busca la cohesión social, tiene identidad cultural (o la construye), es capaz de definir las estrategias de desarrollo concertados entre instituciones locales-regionales y sociedad civil. (Borja , Marzo 2012, pág. 43)

Las ciudades presentan características que permiten identificar la estructura urbana, o al menos la configuración espacial; una de ellas es la función, es decir la manera en que la población ejerce sus fuentes de trabajo predominantes, mismas que cambian de una ciudad a otra según su historia o zona geográfica. La función de una ciudad también representa una forma de clasificar a las ciudades y con ello poder entender la estructura y forma en que se compone; debe recordarse que al plantear la ciudad como un sistema no debiera referirse solo al acomodo de los elementos si no que representa la búsqueda de cómo funcionan éstos en su conjunto.

Es necesario mencionar que, en la búsqueda de clasificar las ciudades por su función, la estadística ha tenido un papel predominante a través de métodos clasificatorios y multivariados que permiten evidenciar y validar las suposiciones descriptivas. Entre las aplicaciones se encuentran la ideada por C.D. Harris en 1943 quien “a partir de la proporción de la población empleada en ciertas profesiones daba una interpretación cuantitativa” (Johnson, 1974, pág. 105) y la clasificaba en etiquetas que fueron ideadas por él, sin embargo es cuestionable que no siempre el mayor porcentaje de personas que trabajan en determinada Actividad Económica, manifiesta una posición en la búsqueda de la clasificación de ciudad con respecto al porcentaje.

El estudio de Harris en las ciudades Estadounidenses, más allá de presentar una clasificación de ciudades, permitió evidenciar grupos de ciudades con porcentajes similares, es decir que las que tenían mayor porcentaje de Actividades Económicas en determinadas áreas como: Comercio al por Mayor, Industriales, Transporte, etc., compartían ubicaciones cercanas entre ellas, lo que generó zonificaciones dentro del país norteamericano.

Otros trabajos que intentaron definir la clasificación de la función en ciudades fueron los propuestos por L. Pownall en 1956 en *The Origins of Towns in New Zeland*, el de H.J. Nelson en Estados Unidos y W. Steigenga en Países Bajos plantearon el uso de la media aritmética como base de la clasificación; el de C.

A. Moser y W. Scott (1961) quienes también realizaron un estudio de carácter estadístico acerca de 157 ciudades de Inglaterra y Gales donde se aplicó la técnica *Componentes Principales* para reducir de 57 variables a 4 componentes o el propuesto por C. Harris (1959) quien determinó, a partir de 30 características del censo de 1959 en 1247 ciudades basado en la misma técnica estadística.

Estos acercamientos a la construcción de modelos en la mitad del siglo XX muestran el intento de definir las relaciones de función con respecto a relaciones espaciales a través de la generación de atributos clasificatorios mismos que, desde el punto de vista metodológico, presentan una herramienta estadística que amalgame el enfoque urbano. En consecuencia, los procesos metodológicos que intentan describir, clasificar y mostrar los comportamientos espaciales de las ciudades toman como vertiente a la estadística para la propuesta de modelos y teorías aplicados a la ciudad.

Algunos estudios, referido a las correlaciones espaciales, han utilizado el Índice de Morán para validar sus aproximaciones; por ejemplo, Moral García (2004) realizó un estudio que muestra la geoestadística como una herramienta para analizar patrones de distribución espacial en trabajos relacionados con el medioambiente y la ecología a partir de muestreos y su aportación permite la realización de mejores interpolaciones en determinados lugares y con esa información se realizan mapas sobre la distribución de diferentes variables.

También Melo & De Freitas (2010) utilizaron esta técnica como indicador de salud para acercarse, a través de un Modelo de Regresión Espacial, un Índice de Necesidades Básicas Insatisfechas en la Región Andina de Colombia, cuyo trabajo consideró variables basadas en el censo poblacional de 2005 y sus resultados mostraron que su modelo representa un buen ajuste a través del valor de proximidad.

Por otra parte, Borrayo & López (2011) se fundamentaron en la Teoría del crecimiento y el desarrollo de nuevas técnicas computacionales en el campo de la economía espacial para generar estudios de análisis de dinámica de la distribución de una medida de ingreso regional que es aplicado en la zona centro de México. Así mismo Barros Díaz & Aroca González (2014) generaron un estudio sobre agrupaciones de envejecimiento en Cuba y sus resultados demostraron la consolidación del proceso de envejecimiento en este país correlacionando sus afectaciones de manera espacial.

Gómez Barroso, Prieto Flores & Moreno Jiménez (2015) realizaron también un estudio ecológico cuyas variables fueron la Razón De Mortalidad Estandarizada, Riesgo Relativo Suavizado y Probabilidad Posterior para identificar los Clúster Espaciales a través del Índice de Morán y cuyos resultados fueron representados cartográficamente.

Finalmente Garrocho, Campos Alanís & Chávez Soto (2018) realizaron un estudio para comparar patrones espaciales de inmuebles colapsados y dañados por los grandes sismos de 1985 y 2017; en dicho estudio los resultados orientaron los esfuerzos de reconstrucción hacia zonas vulnerables y estadísticamente significativas a través del Índice de Morán.

Basado en los trabajos anteriores, se observa que la Autocorrelación espacial, medida a través del Índice de Morán representa una estrategia y herramienta para interpretar los patrones de emplazamiento espacial que genera la distribución de Actividades Económicas en el Territorio y que son medidas a través de los agrupamientos o Clúster.

MATERIALES Y MÉTODOS

La ciudad de Xalapa presenta características económicas puntuales que están directamente relacionadas con su configuración espacial. Para este trabajo se ha tomado como base la variable de Actividad Económica que determina la función de la ciudad, y la Unidad Económica como la observación que se analiza en el modelo. La Clasificación para Actividades Económicas tiene como objetivo “establecer un conjunto jerarquizado de actividades por procesos productivos que pueda ser utilizado para clasificar unidades estadísticas con base en su actividad económica principal” (INEGI, 2010). Enfatizando en esta diferencia, el INEGI plantea que la Actividad Económica “es el conjunto de acciones realizadas por una unidad económica con el propósito de producir o proporcionar bienes y servicios que se intercambian por dinero u otros bienes o servicios”, mientras que la Unidad Económica es “el lugar o entidad donde se realizan las Actividades Económicas”.

Es de evidente que la ciudad de Xalapa no tiene afinidad con ciertas Actividades Económicas, por lo que es necesario reducir el número observaciones de Actividades Económicas para tener un modelo más concreto y significativo.

La base de datos con la que se trabaja en este artículo está sustraída del Directorio Estadístico de Unidades Económicas (DENUE). De forma a priori, se establece una relación sencilla en donde las personas tendrán ocupación siempre y cuando crezca el número de establecimientos económicos que permitan emplear a la población; por el caso contrario si no hubiese la suficiente cantidad de UE, la Población Ocupada (PO) decrecería y aumentaría la Población Desocupada (PD), su complemento.

Para este primer análisis, se ha codificado el listado de Actividades Económicas para simplificar las expresiones estadísticas que permitan observar mejor el comportamiento obteniendo la siguiente tabla:

Tabla 4 Tabla de abreviaciones para análisis Elaboración propia.

Variable	Abreviación
Unidad Económica	UE
Población Ocupada	PO
Población Económicamente Activa	PEA
Actividad Económica	AE
Minería	MI
Servicios Financieros y de Seguros	SFS
Corporativos	CO
Electricidad, Agua y Suministro de Gas por ductos al consumidor final	EAG
Construcción	CT
Industrias Manufactureras	IM
Comercio al por mayor	CMA
Comercio al por menor	CME
Transportes, Correos y Almacenamiento	TCA
Información en Medios Masivos	IMM
Servicios Inmobiliarios y de alquiler de Bienes muebles e intangibles	SIB
Servicios Profesionales, Científicos y Técnicos	SCT
Servicios de apoyo a los negocios y manejo de desechos y servicios de remediación	SND
Servicios Educativos	SE
Servicios de Salud y de asistencia social	SS
Servicios de esparcimiento Culturales y Deportivos, y otros servicios recreativos	SCD
Servicios de alojamiento Temporal y de preparación de Alimentos y bebidas	SAA

Para determinar las Actividades Económicas más significativas que inciden en la Población Ocupada, se plantea un modelo de Regresión Lineal Múltiple bajo la siguiente hipótesis:

$$H_0 : \beta_j = 0 \text{ para } j= 1, 2, 3, \dots, p. (X_j \text{ no influye})$$

vs

$$H_a : \beta_j \neq 0 \text{ para al menos una } j (X_j \text{ si influye})$$

Esta hipótesis se evalúa con la tabla de Análisis de Varianza valorando la R^2 y el p -value asociado a la F_c , así como el análisis de supuestos de homocedasticidad, independencia y normalidad en los residuos.

Una vez obtenido las Actividades Económicas significativas basadas en el modelo de Regresión Lineal Múltiple para determinar la función de la ciudad de Xalapa, se procede a la estimar la correlación espacial, mediante el I de Morán, para determinar las Actividades Económicas que son significativas por su acercamiento a otras en el año 2018 y así evidenciar los patrones de consolidación económica que se presenta en la ciudad.

Lo discutible es ver si estas concentraciones de UE tienen una relación espacial. Para analizar este axioma se utiliza la Autocorrelación espacial del I de Morán que está basada en las ubicaciones y los valores de las entidades simultáneamente. Dado un conjunto de entidades (UE) y un atributo asociado, se evalúa si el patrón expresado está agrupado, disperso o es aleatorio. Bajo este precepto también rompe la estructura

morfológica de la disposición de AGEB y muestra en N km a la redonda la relación que existe entre agrupaciones vecinas, tomando la forma de celdas de manera homogénea. Esto ayuda a explicar el grado en que un objeto es similar a otros objetos cercanos, tomando su importancia para la estadística puesto que ésta se basa en observaciones independientes y dependientes.

RESULTADOS

Al correr el análisis de Regresión en el programa MINITAB 15, se obtiene la siguiente ecuación ajustada de acuerdo al método de mínimos cuadrados que minimiza la suma de los residuos al cuadrado:

Ecuación 1 Ecuación de Regresión Lineal Múltiple que explica las AE que inciden en la PO

$$PO = 0.351 + 0.0042 \text{ SCD} - 0.0193 \text{ TCA} + 0.0070 \text{ SIB} - 0.0353 \text{ IMM} + 0.383 \text{ CME} + 0.0177 \text{ CMA} - 0.0358 \text{ SAA} + 0.117 \text{ IM} + 0.00684 \text{ EAG} + 0.0448 \text{ CT} - 0.0444 \text{ SND} + 0.0668 \text{ SE} + 0.0102 \text{ SS}.$$

Tabla 5 Análisis de variables significativas de la Primera Regresión Lineal Múltiple de PO vs UE

Predictor	Coef.	Coef. de EE	T	P
Constante	0.35133	0.02687	13.08	0.000
SCD	0.00421	0.02587	0.16	0.871
TCA	-0.01931	0.01171	-1.65	0.102
SIB	0.00699	0.02206	0.32	0.752
IMM	-0.03528	0.01569	-2.25	0.026
CME	0.38302	0.04394	8.72	0.000
CMA	0.01770	0.01956	0.90	0.367
SAA	-0.03579	0.04244	-0.84	0.401
IM	0.11651	0.03224	3.61	0.000
EAG	.006843	0.005877	1.16	0.247
CT	0.04479	0.01588	2.82	0.006
SND	-0.04439	0.02381	-1.86	0.065
SE	0.06676	0.03211	2.08	0.040
SS	0.01015	0.01761	0.58	0.565

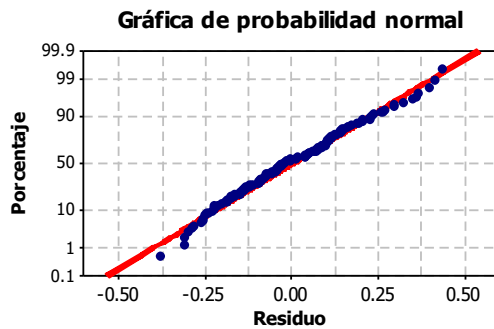
De la tabla anterior se concluye que, con un nivel de confianza del 95% (1- α) y un α de 0.05, los valores de las X de CME, IM, CT, IMM, y SE son estadísticamente significativos para la variable regresora, esto con base en la evidencia del p-value (Sánchez García, Argüello Ortiz, & Montano Rivas, 2018).

Con base en la prueba de Regresión Lineal Múltiple se determinó que las Actividades Económicas más significativas fueron el Comercio al por Menor, la Industria Manufacturera, la Construcción, el Sector Educativo y la Información de Medios Masivos. Las variables antes mencionadas muestran un peso

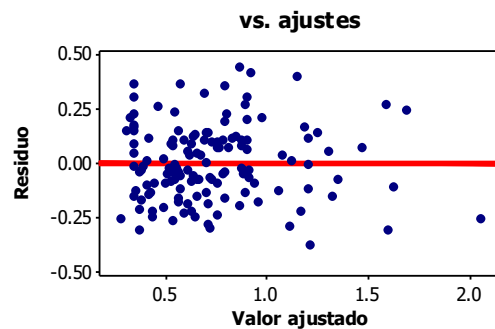
importante en la conformación de las Unidades Económicas dentro de la ciudad, por lo que el siguiente análisis solo se centra en ellas.

Tal como lo refiere (Cuadras, 2007) un estudio más completo del MRLM debe incluir: un análisis gráfico de los residuos, efectos de la colinealidad, mínimos cuadrados ponderados, errores correlacionados, selección de las variables, etc., y para ello debe realizarse el análisis de residuos de manera gráfica como la que se presenta a continuación. Este es un procedimiento que apoya la identificación de estabilidad del análisis del MRLM y sugiere elementos para lograr un mejor ajuste de modelo.

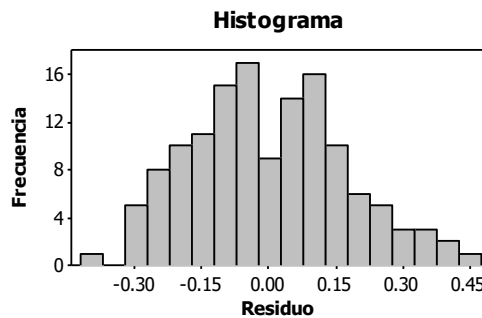
1



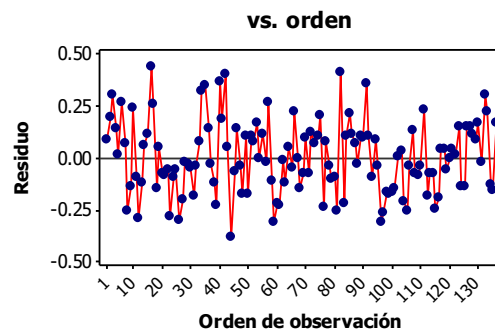
2



3



4



En el gráfico anterior se observa, de manera esquemática y descriptiva, el análisis de los residuos que son la diferencia entre los valores observados y los valores estimados y estos son medidos a partir de la recta. La notación matemática de los residuos es:

Ecuación 2 Notación de Residuos

$$\hat{e}_i = Y_i - \hat{Y}_i$$

Este análisis de residuos aporta elementos significativos para identificar patrones de comportamiento que refuerzan los resultados de las suposiciones teóricas del Análisis de Varianza (ANOVA).

El gráfico 1 (distribución de probabilidad de la normal) representa el supuesto de Normalidad, es decir que los residuos se acercan a la distribución normal $N(0, \sigma^2)$; sin embargo se observa también la presencia de algunos valores atípicos que pueden ser excluidos para determinar si su ausencia permite elevar el Coeficiente de Determinación.

El gráfico 2 (ajustes) representa el supuesto de Homocedasticidad, es decir que la varianza de los errores de la Regresión son los mismos para cada una de las observaciones; este comportamiento va ligado a la Normalidad ya que ésta presenta media 0 y varianza σ^2 constante en todas las variables independientes. Esto lleva a que en el modelo los coeficientes estimados sean los mejores o más eficientes y, de manera gráfica, no se le atribuye ningún patrón de comportamiento que asemeje Heterocedasticidad.

En el gráfico 3 (Histograma) también se presenta, al igual que la número 1, normalidad en los residuos con el primer valor catalogado como dato atípico.

Finalmente en el gráfico 4 se observa el supuesto de Independencia que, bajo la hipótesis de observaciones imparciales, se entiende que los errores son variables aleatorias independientes y se demuestra porque no existe un patrón en el gráfico que represente dependencia. Dicho lo anterior se concluye que al analizar los residuos del modelo, se afirma que se presentan los supuestos teóricos de Normalidad en residuos, Independencia de los errores y Homocedasticidad (homogeneidad de varianza).

El DENUe maneja una base de datos de establecimientos que ingresaron por año al directorio, desde el año 2010 al 2018, por lo que esta periodicidad evidencia de manera puntual hacia qué área geográfica se han ido incrementando las UE perteneciente a los cinco rubros obtenidos. No es factible realizar una representación espacial de los 8 periodos puesto que la cantidad de mapas sería excesiva, por lo que se tomará el 2010 como año de inicio y el 2018 como año final para acrecentar la tendencia al cambio y simplificar su comportamiento.

Una manera de representar esta tendencia al cambio es a través de un mapa temático que represente el grado de concentración de UE; por otra parte una forma de organizar estas concentraciones es mediante la división administrativa de Áreas Geoestadísticas Básicas (AGEB) que estadísticamente muestran una manera fácil de trabajar la base de datos al obtener no más de 200 observaciones.

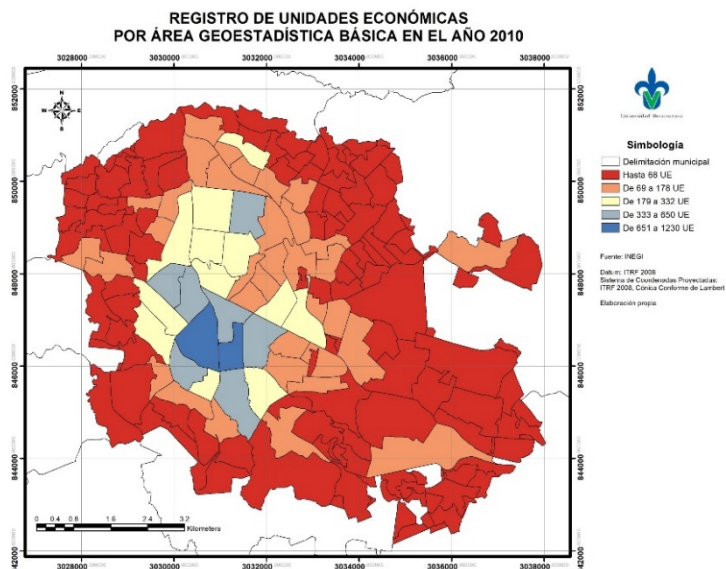


Figura 1 Mapa Temático de la Concentración de Unidades Económicas por Área Geoestadística Básica en el año 2010. Elaboración Propia

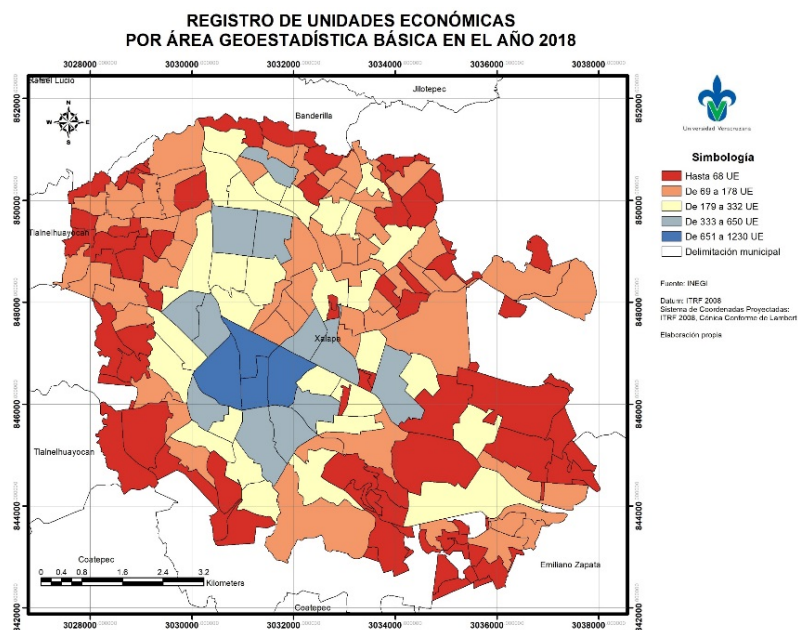


Figura 2 Mapa Temático de la Concentración de Unidades Económicas por Área Geoestadística Básica en el año 2018. Elaboración Propia

En los dos mapas se muestra un cambio por AGEB, tomando un mismo criterio de concentración, que permite identificar un incremento significativo de UE en la zona periférica. En este sentido, organización espacial, entendida como el acomodo de los bloques con diferencia de concentración-dispersión, continúa

manifestando una concentración en el centro administrativo de la ciudad con tendencias al sur y posteriormente se crean anillos marcados hacia la periferia presentando una degradación y una dispersión en dirección a la periferia, recalcando que es el comportamiento de las cinco AE principales en conjunto. Otra manera de representar estos cambios que pudiesen surgir en la fisonomía espacial de la ciudad es a través de una densidad de puntos que muestre la ubicación de las UE en el periodo inicial y los que representan la mitad del periodo hacia el final. Esta contrastación evidencia el lugar geográfico donde se han situado el número de UE y hacia donde presenta un crecimiento o se encuentran bajo el mismo patrón de emplazamiento establecido.

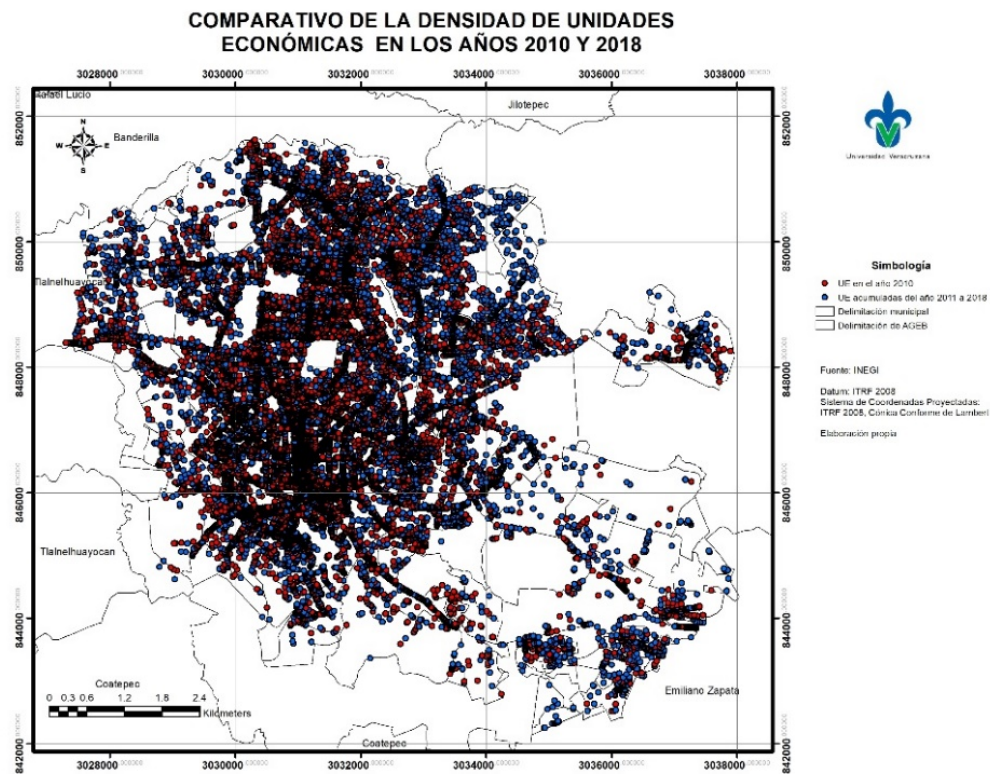


Figura 3 Mapa comparativo de la densidad de Unidades Económicas en los años 2010 y 2018.

Elaboración Propia.

El mapa muestra que la tendencia de incremento de UE hacia la periferia manteniendo el mismo patrón de dispersión, por lo que los comportamientos continúan siendo de la misma organización espacial. Sin embargo, es interesante resaltar que este tipo de representaciones permite observar corredores económicos puntuales, ya que algunas UE muestran su desarrollo espacial tomando como base la estructura vial, por lo que se aprecia una fisonomía en el emplazamiento de las UE con relación a la traza de la ciudad, la cual logra apreciarse a mayor detalle mediante acercamientos regionales.

TRANSICIÓN DE ESTADÍSTICA A CIENCIA DE DATOS

Al aplicar la prueba de Autocorrelación de I de Moran, se identificó de manera espacial, que de las 26 mil 749 UE analizadas hasta el año 2018, sólo el 16.7% resultaron significativas (4,441); es decir, dos mil 886 alcanzaron la clasificación de Clúster y un mil 585 de Valor Atípico, como se parecía en el siguiente mapa.

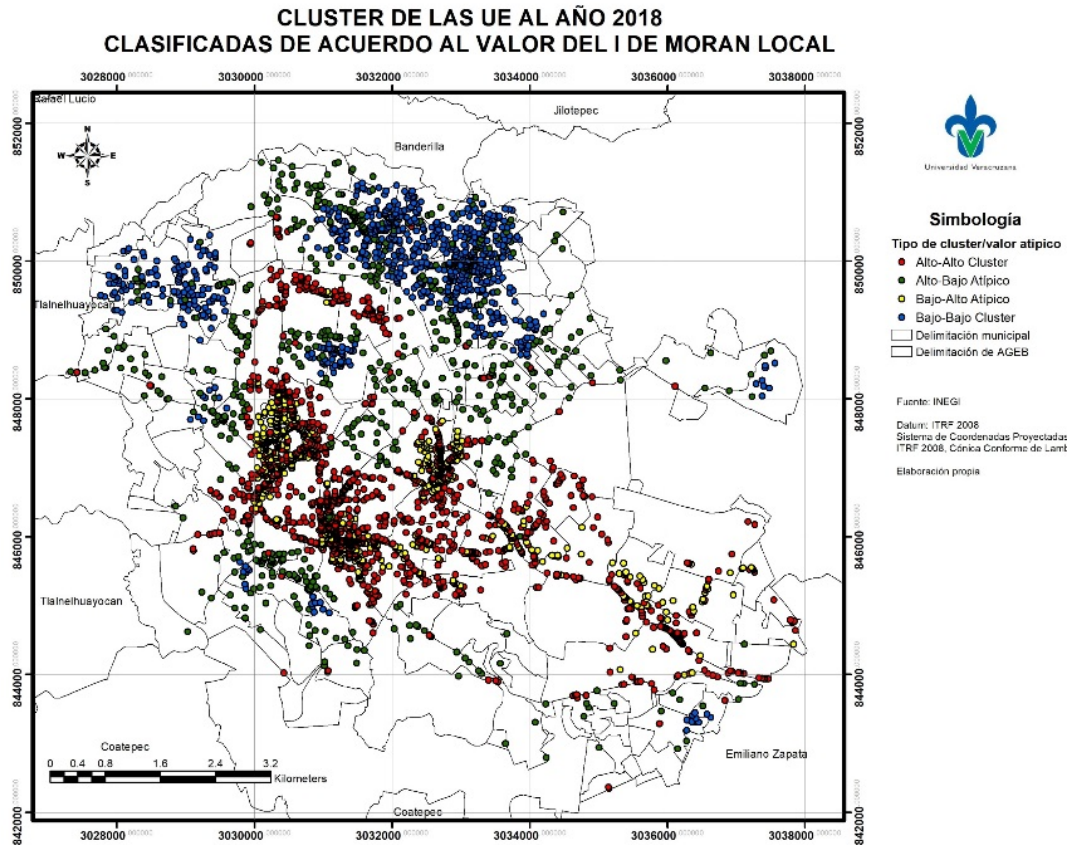


Figura 4 Clúster y Valores atípicos de las Unidades Económicas años al año 2018, clasificadas bajo el I de Moran.

El I de Morán muestra una tipología de anillos no solo en concentración de UE sino en las correlaciones espaciales, es decir que las Zonas que están más correlacionadas son las que se encuentran en el centro administrativo de la Ciudad, es decir en la zona donde se encuentra más consolidadas las UE; sin embargo existen asociaciones que muestran un corredor hacia la zona norte entendiendo que este sector parte de un estrato habitacional popular y que merece especial énfasis; con el apoyo del análisis espacial de concentración de puntos, en un radio de 250 m para las Unidades Económicas, se identificaron concentraciones espaciales estadísticamente significativas entre las dos mil 749 UE, de valores altos (rojos) y bajos (azules), así como valores atípicos altos y bajos, junto con los No Significativos; por lo que un cuestionamiento válido en el tratamiento de esta información es el tamaño de las celdas, ya que al cambiar el tamaño del radio a 500 metros, se aprecia el siguiente resultado.

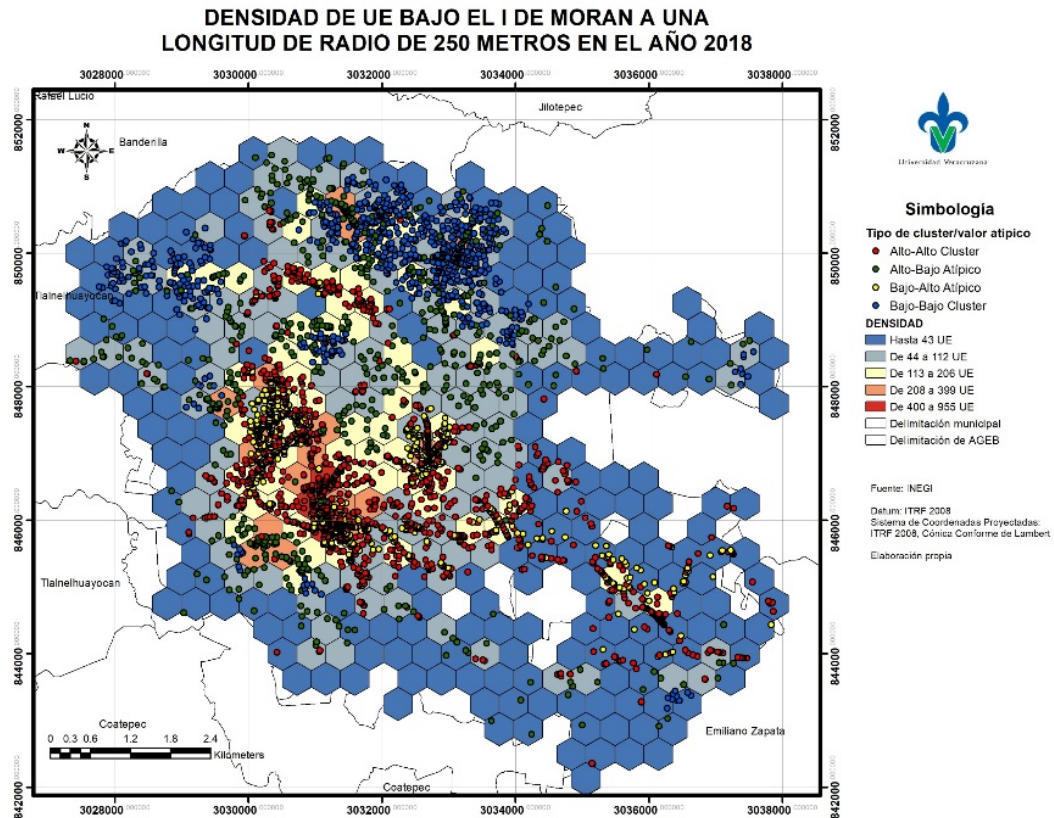


Figura 5 Densidad y Correlación Espacial de Unidades Económicas en una longitud de 250 metros en el año 2018. Elaboración Propia.

Como se muestra en la imagen anterior el análisis de concentraciones presenta un modelo de ciudad donde no solo la mayor concentración de UE se encuentra en el centro administrativo de la ciudad de Xalapa, sino que es en donde existe el mayor grado de asociación. También muestra un grupo de UE con valores atípicos presentados en el norte de la ciudad, donde sólo dos de ellos están asociados pero no muestra una tendencia con las celdas vecinas, es decir existen UE en las zonas continuas pero no están relacionadas estadísticamente, lo que representa los valores No Significativos del I de Moran.

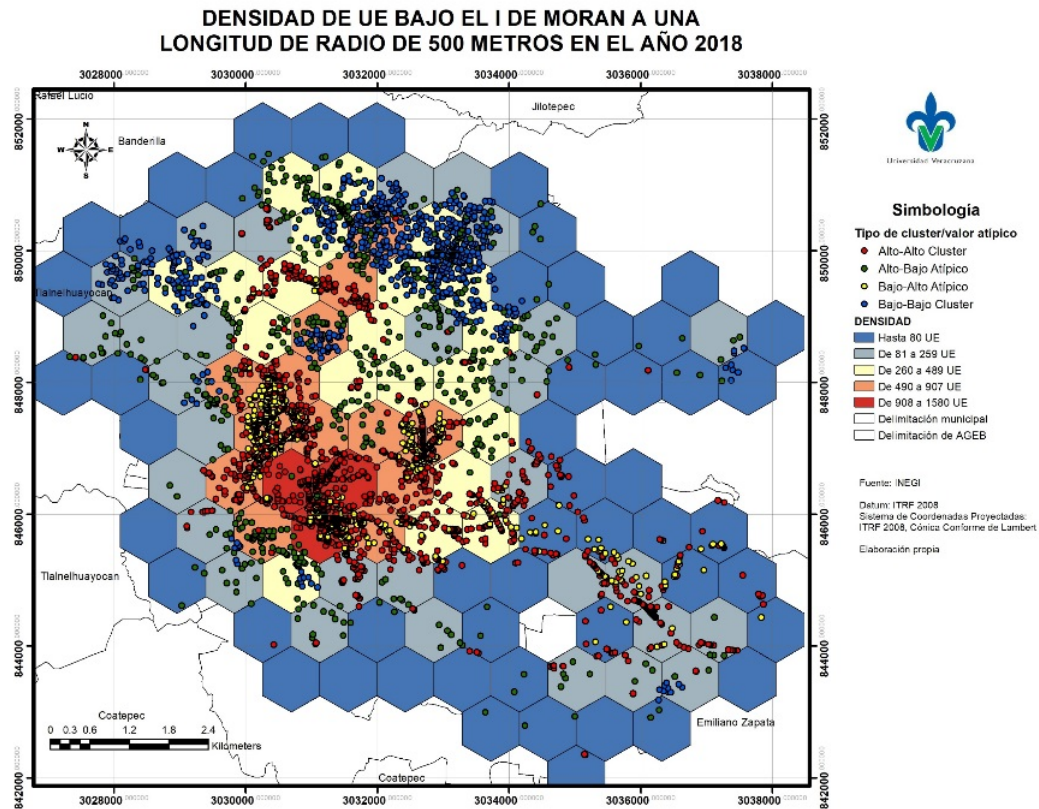


Figura 6 Densidad y Correlación Espacial de Unidades Económicas en una longitud de 500 metros en el año 2018. Elaboración Propia

En este nuevo modelo espacial de concentración se aprecia que existe una asociación fuerte cuando dos o más celdas se integran en una más grande. Esta disposición presenta ya un corredor en consolidación y un centro consolidado con base en agrupaciones de UE, no atendiendo a la estructura vial sino a una organización espacial cuyas piezas homogéneas están fuertemente relacionadas con las celdas vecinas con un nivel de significancia de 0.05.

Una interpretación que es necesario recalcar es que la zona sureste, que limita con el municipio de Emiliano Zapata, se está conformando y está en vías de consolidación como un nuevo núcleo de económico cuya asociación puede llegar a crecer en los próximos años.

Asimismo, este nuevo núcleo pretende llegar a homogeneizar su relación con el centro, por lo que este corredor económico que esta propenso a continuar su expansión produce una consolidación lineal en el marco territorial.

CONCLUSIONES

La ciudad de Xalapa muestra una inclinación hacia Actividades Económicas marcadas al sector terciario, siendo las más significativas, o que más aportan a los trabajos de la población de Xalapa, el Comercio al por Menor, la Industria Manufacturera, la Construcción, el Sector Educativo y la Información de Medios Masivos.

Estas Actividades Económicas están representadas en Unidades Económicas o establecimientos laborales de una manera heterogénea, cuyo comportamiento más fuerte está en el centro administrativo de la ciudad basado en la densidad de UE dentro del territorio. También se observan cambios significativos en la concentración de UE dentro del territorio en un periodo de 2010 a 2018, que reflejan que a medida que crecen en un sector también lo hacen en otro sector. El mapa temático muestra una variación estadística de concentración, es decir que no tiene cambio por lo que cambia a un sistema heterogéneo. Dado este cambio es oportuno recurrir a herramientas estadísticas que expliquen espacialmente si esta dispersión y configuración espacial muestra una autocorrelación, por lo que el I de Moran es una técnica usada para validar este principio y representarlo.

El I de Morán muestra una organización tal que las UE centrales tienen mayor correlación con las UE vecinas con características en común y este patrón asociativo se extiende en un corredor de desarrollo económico muy marcado hacia el norte solo en las AE que se colocaron en el modelo, lo que permite observar un eje consolidación económica.

El trabajo solo evidenció el diálogo entre disciplinas, los procedimientos e interpretaciones de disciplinas generan un criterio distinto en la búsqueda de nuevo conocimiento o formas de explicar los fenómenos que en un momento concentran limitaciones metodológicas o que los primeros resultados no arrojan una conclusión significativa.

Se considera que cada AE significativa presenta un comportamiento espacial diferente por lo que cada uno de ellos manifestaría una composición característica que permitiría evidenciar los sectores económicos relacionados con el territorio de la ciudad.

Conceptualmente, y basado en la variable económica, se aprecia un modelo de ciudad centralizado donde los ajustes a las políticas públicas y los planes de Desarrollo Urbano pueden estar sujetos a descentralizar la concentración económica y homogeneizar zonas que quizás carecen de estas Actividades Económicas, lo cual repercutiría en traslados de trabajo o de abastecimiento y en la consolidación de otras zonas de la ciudad.

La herramienta del I de Morán muestra que, al inhibir el límite administrativo de las AGEB, contribuye a la generación de celdas homogéneas que permite una asociación de correlación espacial, evidenciando que lo importante es la interpretación de los modelos desde la perspectiva urbana.

Este tipo de trabajos no permite establecer una sola conclusión determinista ya que cada paso, técnica y enfoque manifiesta una propia aportación al aspecto metodológico, epistemológico e incluso holístico del trabajo de investigación. Por esta razón cada segmento de este trabajo genera una exploración a las bondades y deficiencias que se pueden encontrar antes, durante y al final de este proceso.

Una de las vertientes que hoy en día genera las buenas aportaciones al campo del conocimiento es reconocer que la disciplina por si sola carece de técnicas, instrumentos, conceptos e incluso metodologías para explicar fenómenos que cuestionen paradigmas. Por ello, pensando en hacia dónde va la nueva investigación se abren brechas hacia el diálogo multidisciplinar que, desde el punto de vista del autor, no es más que tomar conceptos o metodologías de otra disciplina para explicar lo que la disciplina propia no puede realizar, haciendo un balance quizás entre las analogías para obtener una mejor interpretación desde el urbanismo.

El diálogo con otras disciplinas no funge en este trabajo como punto central sino que al realizarse éste, es posible sustentar y argumentar conceptos, constructos y axiomas que el urbanismo reconoce poco o que no se aventura hibridar nuevas metodologías contemporáneas quedando en un urbanismo clásico y hasta cierto punto repetitivo y predecible.

Las herramientas digitales, también fungen un papel importante ya que facilitan y sintetizan los procesos de análisis dentro de un trabajo de investigación, por lo que la carencia de ellos puede generar sesgos y tiempo de trabajo innecesario. Cabe destacar que este tipo de herramientas y técnicas no solo se presentan en el diseño de gráficos sino en la validación de instrumentos que permiten generar certeza de lo que se propone o se interpreta.

Finalmente, la búsqueda de metodologías prestadas de otras disciplinas no demerita el eje rector sobre el cual se hace investigación, sino que aporta una visión, validación y dimensión del contenido presentando una manera distinta de explicar el mismo fenómeno pero interpretado desde la visión urbanística, por lo que la asociación de frontera y tocar puntos en común mediante los cuales se puede evidenciar de una manera adecuada el fenómeno.

REFERENCIAS

- Aponte Gómez , C., Romero Aroca, E. M., & Santa Guzmán, L. F. (2015). Análisis de datos espaciales del Índice de Necesidades Básicas Insatisfechas en la Región Andina. *Perspectiva Geográfica*, 391-448.
- Barros Díaz, O., & Aroca Gonzalez , P. (2014). Econometría espacial y el análisis sociodemográfico. Aplicación en la formación de agrupaciones espaciales de envejecimiento en Cuba, período 2003-2009. *Novedades en población*.
- Borja , J. (Marzo 2012). *Revolución urbana y derechos ciudadanos: Claves para interpretar las contradicciones de la ciudad actual*. Barcelona, España: Departamento de Geografía Humana: Facultad de Geografía e Historia de la Universidad de Barcelona.
- Borrayo Lopez, R., & Castañeda Arriaga, J. (2011). Análisis de transición dinámica: un enfoque no paramétrico aplicado a la región centro de México (1988-2003). *Revista Problemas del Desarrollo*.
- Bustelo, P. (1998). *Teorías Contemporáneas del Desarrollo Económico*. Madrid: Síntesis.
- Garrocho, C., Campos Alanís, J., & Chávez Soto, T. (2018). Análisis espacial de los inmuebles dañados por el sismo 19S-2017 en la Ciudad de México. *salud pública de méxico*.
- Gomez Barroso, D., Prieto Flores, M. E., & Moreno Jimenez, A. (2015). Análisis espacial de la mortalidad por enfermedades cardiovasculares en la ciudad de Madrid, España. *Revista Española de Salud Pública*, 27-37.
- INEGI. (2010). *Clasificación para Actividades Económicas*.
- Johnson, J. (1974). *Geografía Urbana*. España: Oikos_tau.
- Melo, E. C., & De Freitas Mathias, T. A. (2010). Distribución y auto-correlación espacial de indicadores de la salud de la mujer y del niño en el estado de Paraná, Brasil. *Latino-Am. Enfermagem*.
- Moral García, F. J. (2004). Aplicación de la geoestadística en las ciencias ambientales . *Ecosistemas*, 78-86.
- Sánchez García, J., Argüello Ortiz, A., & Montano Rivas, J. (2018). Análisis Geoestadístico de la Estructura Urbana de Xalapa en el periodo 2010. En *Aplicaciones de Metodología Estadística* (págs. 10-39). Xalapa: Imaginaria Editores. doi:978-84-685-2102-2
- W. Richarson, H. (1973). *Economía Regional; Teoría de la localización, estructuras urbanas y crecimiento regional*. Barcelona: Vicens-Vives.

APLICACIÓN DE SOFTWARE LIBRE EN LA MINERÍA DE DATOS: R Y PYTHON COMO HERRAMIENTAS DE ANÁLISIS

Alexis Iván Mora Argüello¹, Cecilia Cruz López¹, Sergio Francisco Juárez Cerrillo¹,
Ernesto Pedro Menéndez Acuña², José Juan Muñoz León¹

¹Facultad de Estadística e Informática, ²Facultad de Matemáticas

Universidad Veracruzana

e-mail: alexis.mora.ar@hotmail.com, ceccruz@uv.mx, sejuarez@uv.mx, emenendez@uv.mx; juanmunoz@uv.mx

RESUMEN

El capítulo presenta los resultados de la aplicación de técnicas de minería de datos utilizando los lenguajes R y Python, con el propósito de comparar su desempeño, precisión y facilidad de uso. Se implementaron cuatro métodos: vecinos más cercanos (k-NN), árboles de decisión, bosques aleatorios y análisis de *clusters*, aplicados a una base de datos de diamantes del Gemological Institute of America (GIA).

Los resultados mostraron que ambos entornos alcanzaron niveles de precisión similares en los modelos supervisados, con valores superiores al 80%. Python destacó por su flexibilidad, amplia gama de librerías especializadas y capacidad para generar visualizaciones gráficas de alta calidad. Por su parte, R ofreció una estructura más directa para la modelación estadística y un proceso de estandarización más simple, lo que facilita su uso en entornos académicos.

En los análisis no supervisados, como el *clustering*, ambos softwares identificaron patrones consistentes; sin embargo, Python logró una mejor definición de los grupos gracias a su manejo de componentes principales. En conjunto, los hallazgos evidencian que R y Python son herramientas complementarias que potencian el análisis de grandes volúmenes de información y favorecen la formación de competencias en ciencia de datos y estadística aplicada.

Palabras clave: Minería de datos; software libre; R; Python; clasificación; aprendizaje automático; árboles de decisión

INTRODUCCIÓN

En la actualidad, la producción masiva de datos en el mundo ha adquirido proporciones sin precedentes, impulsada por el desarrollo de tecnologías digitales, plataformas en línea y la expansión del Internet de las cosas. Cada segundo se generan cantidades inmensas de información que, en muchos casos, quedan

sin analizar, lo que evidencia la necesidad creciente de herramientas capaces de procesar, depurar y extraer conocimiento útil de dichos volúmenes de datos. Moreno (2018) señala que en 2018 la cantidad de datos generados alcanzó los 33 zettabytes, es decir, 16.5 veces más que nueve años atrás, lo que demuestra la velocidad exponencial con que crece la información global.

En este contexto, la minería de datos surge como una metodología fundamental que permite descubrir patrones, relaciones y tendencias ocultas en grandes bases de datos. Según Han, Pei y Kamber (2011), se trata de un proceso mediante el cual se aplica inteligencia analítica para transformar datos en conocimiento. En México, Aguado (2018) destaca su uso en sectores estratégicos como el comercio, la banca y los servicios públicos, donde permite segmentar clientes, optimizar procesos y predecir comportamientos. Asimismo, instituciones gubernamentales y empresas como CFE, PEMEX, TELMEX, IMSS o ISSSTE recurren cada vez más a la analítica masiva de datos como apoyo para la toma de decisiones.

En el ámbito académico, la Universidad Veracruzana ha reconocido la importancia de esta disciplina al incluir en sus programas educativos asignaturas como 'Minería de datos y aprendizaje máquina' y 'Estadística multivariada', fortaleciendo la formación de profesionales capaces de utilizar herramientas como R y Python, considerados los softwares de libre acceso más poderosos en la actualidad para el análisis estadístico y la minería de datos.

El propósito de este capítulo es describir y ejemplificar el uso de R y Python en el proceso de análisis de minería de datos, comparando sus características, ventajas y requerimientos técnicos. Ambos programas ofrecen soluciones robustas, gratuitas y versátiles para el manejo de grandes volúmenes de información, lo que los convierte en alternativas esenciales frente a softwares comerciales como SPSS o MINITAB (Carvajal y Maita, 2008; Miranda, 2015).

R, desarrollado en los años noventa por Ross Ihaka y Robert Gentleman, es un lenguaje de programación estadístico de código abierto que hereda las capacidades de S-Plus y que ha ganado popularidad entre analistas y científicos de datos por su sintaxis flexible, su gran comunidad de usuarios y sus potentes librerías gráficas (Gatto, 2018). Por su parte, Python, creado por Guido van Rossum en 1989, combina potencia computacional con una sintaxis clara y legible, siendo ampliamente utilizado en ciencia de datos, inteligencia artificial y minería de datos mediante bibliotecas como Pandas, NumPy, Matplotlib y PySpark (Montoro, 2012; Python, 2020).

Ambos entornos permiten implementar técnicas estadísticas fundamentales en minería de datos, como árboles de decisión (Berlanga, Rubio y Vilà, 2013), bosques aleatorios (Medina y Ñique, 2017), algoritmos de vecinos más cercanos (IBM, 2008) y métodos de agrupamiento como el clustering (Figueras, 2001).

Estas técnicas, integradas en los lenguajes R y Python, ofrecen al analista un marco integral para la exploración, modelado y visualización de datos complejos.

El auge de estas herramientas refleja la tendencia global hacia la adopción de software libre, que permite democratizar el acceso a la analítica avanzada sin depender de licencias costosas. Este trabajo busca también contribuir a la enseñanza de estas metodologías, sirviendo como guía práctica para estudiantes y profesionales interesados en comprender las capacidades comparativas de R y Python en el campo de la minería de datos. Finalmente, la integración de ambas herramientas en proyectos de investigación, como el estudio de Lenton (2013) sobre votaciones legislativas en Argentina, demuestra su potencial combinado para la adquisición, limpieza, análisis y visualización de información compleja.

En síntesis, la presente INTRODUCCIÓN ofrece un panorama sobre la relevancia de la minería de datos en la era digital, resaltando el papel estratégico de R y Python como instrumentos de análisis accesibles, versátiles y en constante evolución. En la siguiente sección se detalla la metodología empleada para ejemplificar la aplicación práctica de ambos softwares en el proceso de minería de datos.

MATERIALES Y MÉTODOS

El presente estudio se desarrolló bajo un enfoque retrospectivo, dado que los datos analizados fueron obtenidos a partir de registros secundarios, sin intervención directa en su recolección. Este tipo de diseño implica que los investigadores no tuvieron participación en la obtención de las mediciones originales, por lo que no es posible garantizar la exactitud con la que fueron realizadas. Asimismo, el diseño se considera explicativo, ya que busca identificar y describir el comportamiento de una variable en función de otra, siguiendo un enfoque causal que exige control y cumplimiento de los criterios básicos de causalidad. En este sentido, el propósito metodológico consistió en analizar las relaciones existentes entre las características morfológicas de los diamantes y sus atributos de clasificación mediante técnicas de minería de datos implementadas en los lenguajes R y Python.

La información utilizada proviene de la base de datos disponible en la página oficial del Gemological Institute of America (GIA), la cual constituye una fuente de referencia internacional en la certificación de piedras preciosas. La población de estudio estuvo conformada por 53,940 registros de diamantes, cada uno descrito por diez variables que representan diferentes propiedades físicas y ópticas. Con el fin de realizar un análisis más focalizado, se seleccionó una muestra compuesta por tres de las variables más representativas: table, profundidad y corte, consideradas esenciales para la clasificación y valoración de

TRANSICIÓN DE ESTADÍSTICA A CIENCIA DE DATOS

los diamantes. Estas variables se eligieron por su relevancia en la determinación de la calidad y apariencia del diamante, así como por su utilidad para la implementación de algoritmos de clasificación.

El proceso de operacionalización de variables permitió definir de manera precisa los conceptos analizados, sus equivalentes operativos, el tipo de variable y la escala de medición correspondiente. Entre las variables cuantitativas continuas incluidas destacan el quilate, la profundidad, table y las dimensiones X, Y y Z, todas medidas en escala de razón, lo que permite comparaciones proporcionales entre valores. Las variables cualitativas comprenden corte, color y claridad, las cuales se manejaron en escalas ordinal y nominal según el tipo de clasificación. El precio del diamante se consideró una variable cuantitativa discreta expresada en euros. Este proceso de operacionalización permitió estructurar el análisis estadístico conforme a los supuestos requeridos por las técnicas de minería de datos seleccionadas (Ver Tabla 1).

Tabla 1. Operacionalización de las variables

VARIABLE	DEFINICIÓN CONCEPTUAL	DEFINICIÓN OPERACIONAL	TIPO DE VARIABLE	ESCALA DE MEDICIÓN
Quilate	Unidad de masa de las perlas y piedras preciosas.	Unidad básica de medida de peso del diamante.	Cuantitativa Continua.	Razón
Corte	Forma de un diamante que es lo que determina en última instancia su capacidad de brillar.	Clasificación que se le da al tipo de corte del diamante.	Cualitativa Dicotómica.	Ordinal
Color	Impresión producida por un tono de luz en los órganos visuales.	Clasificación que se le da al tipo de diamante de acuerdo con su falta de color.	Cualitativa Politómica.	Nominal
Claridad	Abundancia de luz	Clasificación que se le da al diamante se refiere a la ausencia de inclusiones y defectos.	Cualitativa Politómica.	Ordinal
Profundidad	Distancia entre el fondo de algo y el punto tomado como referencia (parte más alta, entrada, borde, etc.)	Es el porcentaje que se le da con relación a la altura del diamante dividida entre por la anchura.	Cuantitativa Continua.	Razón
Table	El labrado por la haz con una superficie plana y alrededor con cuatro biseles.	Es el porcentaje que mide la faceta superior y la más grande.	Cuantitativa Continua.	Razón
Precio	Cantidad de dinero que permite la adquisición o uso de un bien o servicio.	Precio designado en unidad euro al diamante.	Cuantitativas Discreta	Razón

TRANSICIÓN DE ESTADÍSTICA A CIENCIA DE DATOS

X	Signo que se emplea en lugar de un nombre que no se quiere o no se puede decir.	Raza de diamantes técnicamente blancos que a menudo se pasa por alto.	Cuantitativa Continua.	Razón
Y	Signo que representa la segunda incógnita de una ecuación o un sistema de ecuaciones.	Raza de diamantes técnicamente blancos que a menudo se pasa por alto.	Cuantitativa Continua.	Razón
Z	Signo que representa la tercera incógnita de una ecuación o un sistema de ecuaciones.	Raza de diamantes técnicamente blancos que a menudo se pasa por alto.	Cuantitativa Continua.	Razón

El análisis estadístico se fundamentó en la aplicación de técnicas de minería de datos orientadas a la identificación de patrones y la clasificación de los registros mediante distintos algoritmos. Entre los métodos utilizados se incluyeron los árboles de decisión, bosques aleatorios, el algoritmo del vecino más cercano (k-NN) y el método de agrupamiento (clustering). Estas técnicas se implementaron de manera comparativa en los entornos de programación R y Python, a fin de contrastar la eficiencia, precisión y calidad de las salidas gráficas generadas por cada software.

El desarrollo metodológico se alineó con los objetivos de la investigación. El primer objetivo, enfocado en describir las características de ambos tipos de software en la minería de datos, se alcanzó tras la ejecución completa de los análisis y la evaluación de las herramientas disponibles en cada entorno. El segundo objetivo, orientado a determinar las ventajas y desventajas de R y Python en el procesamiento y modelado de datos, se logró mediante la observación del rendimiento y la facilidad de uso durante el desarrollo del estudio. El tercer objetivo, relacionado con la identificación de los requerimientos básicos de instalación y operación de ambos programas, se abordó antes del análisis, documentando los aspectos técnicos esenciales para su implementación. Finalmente, el cuarto objetivo, centrado en formular argumentos sobre la calidad y representatividad de las salidas gráficas, se cumplió tras la interpretación visual de los resultados y la comparación de las capacidades de visualización de cada software.

En conjunto, los materiales y métodos empleados proporcionan una estructura metodológica sólida que combina el rigor estadístico con la flexibilidad de las herramientas de software libre, permitiendo un análisis reproducible y transparente de grandes volúmenes de información. La implementación paralela de R y Python permitió contrastar su desempeño, evidenciando el potencial de ambas plataformas para la enseñanza y la práctica profesional en minería de datos.

RESULTADOS

El presente capítulo muestra los resultados derivados de la aplicación de diversas técnicas de minería de datos mediante los softwares R y Python. Se presentan los hallazgos obtenidos a partir de los algoritmos de vecinos más cercanos (k-NN), árboles de decisión, bosques aleatorios y análisis de clusters, enfatizando la comparación entre ambos entornos de programación en términos de precisión, facilidad de implementación y capacidad gráfica. Para mantener la claridad y la relevancia visual, se seleccionaron únicamente las figuras más representativas de cada procedimiento.

El primer análisis corresponde al algoritmo de vecinos más cercanos (k-NN) implementado en Python. En este entorno, se realizó la carga de los paquetes y la lectura del conjunto de datos (Figura 1), seguido de la separación de los datos en conjuntos de entrenamiento y prueba. El modelo configurado con la métrica de Minkowski ($p=2$), equivalente a la distancia euclidiana, alcanzó una precisión del 82%, según la matriz de confusión (Figura 2). Este resultado indica un desempeño satisfactorio del algoritmo en la clasificación del tipo de corte de los diamantes.

```
In [1]: import pandas as pd
import numpy as np
import sklearn
from sklearn.model_selection import train_test_split
```

```
In [2]: dataset = pd.read_csv('C:\\Users\\alexi\\OneDrive\\Escritorio\\diamantesDosClas.csv')
dataset.head()
```

```
Out[2]:
```

	quilate	corte	color	claridad	profundidad	table	precio	x	y	z
0	0.23	Good	E	SI2	61.5	55.0	326	3.95	3.98	2.43
1	0.21	Premium	E	SI1	59.8	61.0	326	3.89	3.84	2.31
2	0.23	Good	E	VS1	58.9	65.0	327	4.05	4.07	2.31
3	0.29	Premium	I	VS2	62.4	58.0	334	4.20	4.23	2.63
4	0.31	Good	J	SI2	63.3	58.0	335	4.34	4.35	2.75

Figura 1. Carga de paquetes y lectura de datos

```
In [10]: from sklearn.metrics import confusion_matrix
```

```
In [11]: cm = confusion_matrix(y_test, y_pred)
cm
```

```
Out[11]: array([[5895, 1058],
               [1265, 5267]], dtype=int64)
```

```
In [12]: classifier.score(X_test, y_test)
```

```
Out[12]: 0.8277345198368558
```

Figura 2. Matriz de confusión.

TRANSICIÓN DE ESTADÍSTICA A CIENCIA DE DATOS

En R, el mismo algoritmo fue replicado empleando las librerías 'class' y 'caret'. La Figura 3 muestra el proceso de estandarización de los datos, paso esencial para evitar discrepancias entre las escalas de las variables independientes. Posteriormente, se realizó la partición de los datos para entrenamiento y prueba, alcanzando una precisión del 83% (Figura 4). Estos resultados sugieren que tanto R como Python presentan un comportamiento similar en términos de desempeño predictivo, aunque R ofrece mayor sencillez en la estandarización y evaluación del modelo.

```
> datos$profundidad.z = scale(datos$profundidad)
> datos$table.z = scale(datos$table)
> head(datos)
```

	quilate	corte	color	claridad	profundidad	table	precio	x	y	z
1	0.23	Good	E	SI2	61.5	55	326	3.95	3.98	2.43
2	0.21	Premium	E	SI1	59.8	61	326	3.89	3.84	2.31
3	0.23	Good	E	VS1	56.9	65	327	4.05	4.07	2.31
4	0.29	Premium	I	VS2	62.4	58	334	4.20	4.23	2.63
5	0.31	Good	J	SI2	63.3	58	335	4.34	4.35	2.75
6	0.24	Premium	J	VVS2	62.8	57	336	3.94	3.96	2.48

```

profundidad.z    table.z
1    -0.1740899 -1.0996618
2    -1.3607259  1.5855140
3    -3.3849872  3.3756312
4     0.4541292  0.2429261
5     1.0823482  0.2429261
6     0.7333376 -0.2046032
```

Figura 3. Estandarización de datos.

```
> pred2 = knn(train[,11:12], test[,11:12], train[,2], k=7)
> errmat2 = table(test$corte, pred2, dnn = c ("Actual", "Predichos"))
> errmat2
```

	Predichos	
Actual	Good	Premium
Good	6007	1009
Premium	1198	5270

```
> precisión <- (sum(diag(errmat2)))/sum(errmat2)
> precisión
[1] 0.8363245
```

Figura 4. Matriz y precisión

En cuanto a los árboles de decisión, Python mostró un rendimiento estable al emplear el algoritmo 'DecisionTreeClassifier'. La evaluación de los datos de entrenamiento y prueba alcanzó valores de precisión entre 83% y 84% (Figura 5). La visualización del árbol generado (Figura 6) permitió identificar reglas claras para clasificar los diamantes según la profundidad y el porcentaje de tabla, proporcionando una estructura lógica y jerárquica del proceso de decisión.

```

In [13]: arbol=DecisionTreeClassifier()

In [14]: arbol.fit(x_train, y_train)

Out[14]: DecisionTreeClassifier(class_weight=None, criterion='gini', max_depth=None,
max_features=None, max_leaf_nodes=None,
min_impurity_decrease=0.0, min_impurity_split=None,
min_samples_leaf=1, min_samples_split=2,
min_weight_fraction_leaf=0.0, presort=False, random_state=None,
splitter='best')

In [15]: arbol.score(x_test, y_test)

Out[15]: 0.8367939686070943

In [16]: arbol.score(x_train, y_train)

Out[16]: 0.8432650034429789

```

Figura 5. Aplicación del algoritmo.

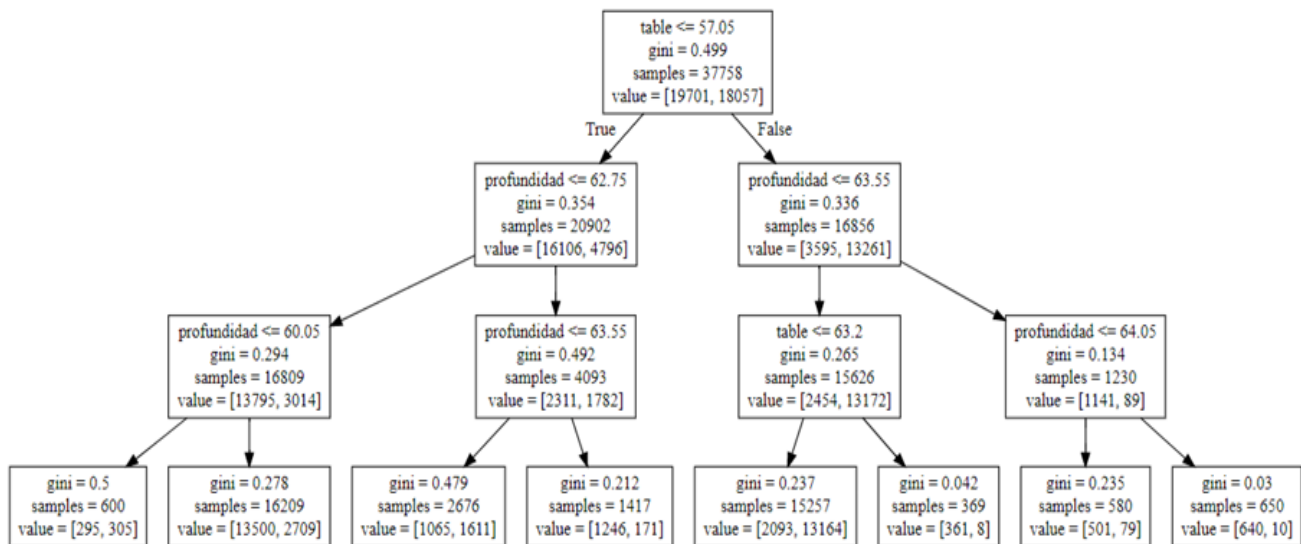


Figura 6. Árbol de decisión Python.

En R, el árbol de decisión se modeló utilizando la librería 'rpart', y la visualización se efectuó mediante 'rpart.plot'. El Figura 7 muestra las divisiones principales del modelo: si el porcentaje de tabla es menor a 57, el diamante se clasifica como 'Good', mientras que valores superiores a este umbral se asocian con cortes 'Premium'. Esta representación facilita la interpretación del modelo y la comparación con los resultados obtenidos en Python.

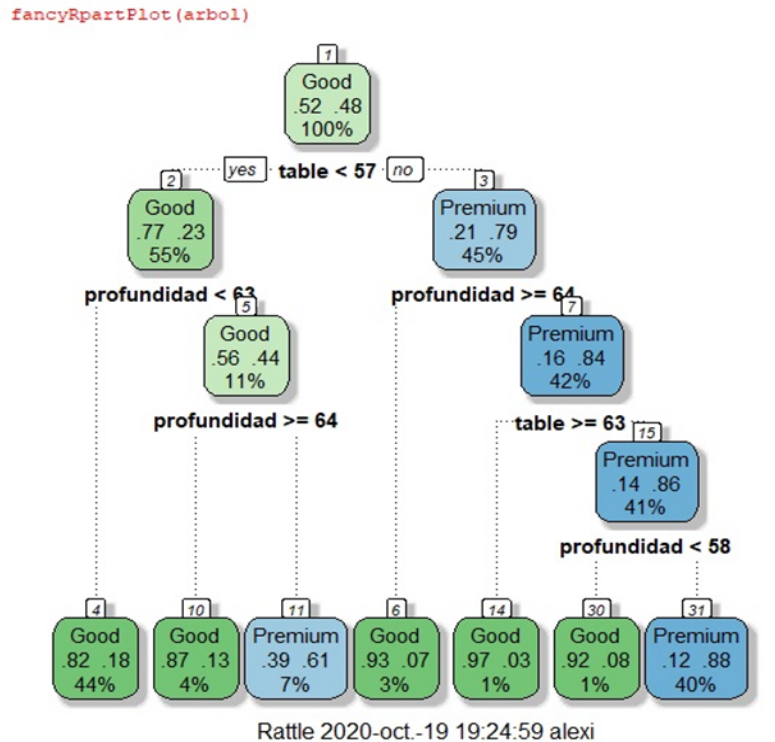


Figura 7. Árbol de decisión R.

En la aplicación de bosques aleatorios, Python utilizó el algoritmo 'RandomForestClassifier' de la librería Sklearn (Figura 8). Los resultados mostraron una precisión del 84%, de acuerdo con la matriz de confusión (Figura 9). El modelo logró identificar correctamente la mayoría de los casos de corte 'Good' y 'Premium', lo que refleja su capacidad para manejar la variabilidad en los datos. En R, mediante las librerías 'caret' y 'randomforest', el modelo alcanzó resultados comparables, con un total de 7165 casos 'Good' y 6371 'Premium' correctamente clasificados (Figura 1).

```

In [6]: modelo = RandomForestClassifier(
        random_state      = 1,
        n_estimators      = 666,
        min_samples_split = 2,
        min_samples_leaf  = 1,
        n_jobs            = 1
      )
        modelo.fit(X = X_train, y = y_train)

Out[6]: RandomForestClassifier(bootstrap=True, class_weight=None, criterion='gini',
        max_depth=None, max_features='auto', max_leaf_nodes=None,
        min_impurity_decrease=0.0, min_impurity_split=None,
        min_samples_leaf=1, min_samples_split=2,
        min_weight_fraction_leaf=0.0, n_estimators=666, n_jobs=1,
        oob_score=False, random_state=1, verbose=0, warm_start=False)

In [7]: prediccion = modelo.predict(X_test)
  
```

Figura 8. Aplicación del modelo y predicción.

```
In [8]: print(metrics.classification_report(y_true=y_test, y_pred=prediccion))
```

	precision	recall	f1-score	support
Good	0.84	0.85	0.84	6940
Premium	0.84	0.82	0.83	6545
micro avg	0.84	0.84	0.84	13485
macro avg	0.84	0.84	0.84	13485
weighted avg	0.84	0.84	0.84	13485

```
In [9]: print(pd.crosstab(y_test, prediccion, rownames=['REAL'], colnames=['PREDICCION']))
```

PREDICCION	Good	Premium
REAL		
Good	5903	1037
Premium	1149	5396

Figura 9. Matriz de confusión.

```
> pred = predict(mod, datos[-training.ids,])
> table(datos[-training.ids,"corte"], pred, dnn=c ("actual", "predicho"))
      predicho
actual   Good Premium
Good    7165    1255
Premium 1390    6371
```

Figura 10. Predicción R.

Finalmente, en el análisis de clusters se aplicó el método k-means. En Python, la técnica del código de Jambú permitió determinar el número óptimo de clusters en tres grupos (Figura 11). Posteriormente, mediante la reducción de dimensionalidad con componentes principales (PCA), se visualizó una clara separación entre los grupos (Figura 12), lo que confirma la consistencia del agrupamiento.

```
In [7]: wcss = []  
  
for i in range(1, 11):  
    kmeans = KMeans(n_clusters=i, max_iter = 300)  
    kmeans.fit(dataset_norm)  
    wcss.append(kmeans.inertia_)
```

```
In [8]: plt.plot(range(1, 11), wcss)  
plt.title ("Codo de Jambu")  
plt.xlabel ("Numero de Clusters")  
plt.ylabel("WCSS")  
plt.show()
```

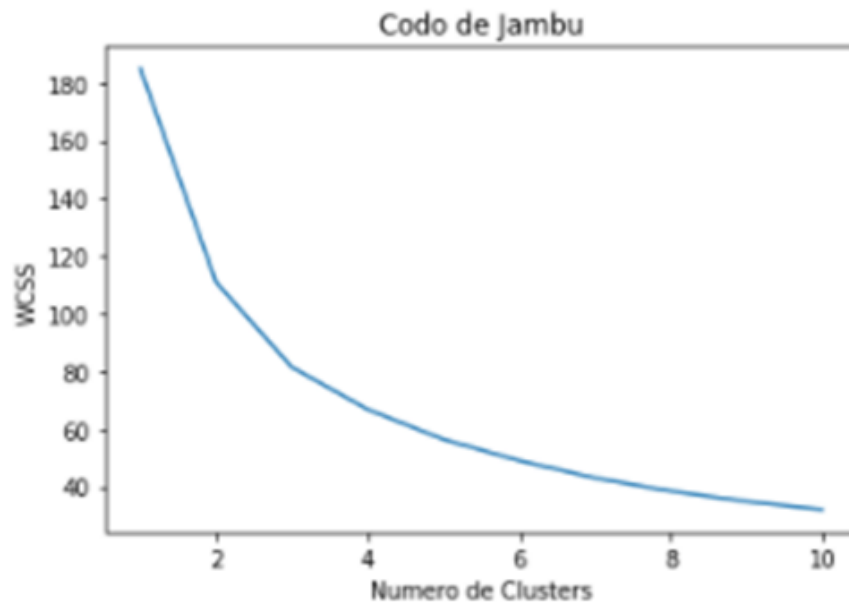


Figura 11. Codo de Jambú Python.

```
In [12]: fig = plt.figure(figsize = (6,6))

ax= fig.add_subplot (1,1,1)
ax.set_xlabel("Componente 1", fontsize = 15)
ax.set_ylabel("Componente 2", fontsize = 15)
ax.set_title("Componentes Principales", fontsize=20)

color_theme = np.array(["blue", "green", "orange"])
ax.scatter(x = pca_nombres_dataset.Componente_1, y = pca_nombres_dataset.Componente_2,
          c=color_theme[pca_nombres_dataset.KMeans_Clusters], s= 50)
plt.show()
```

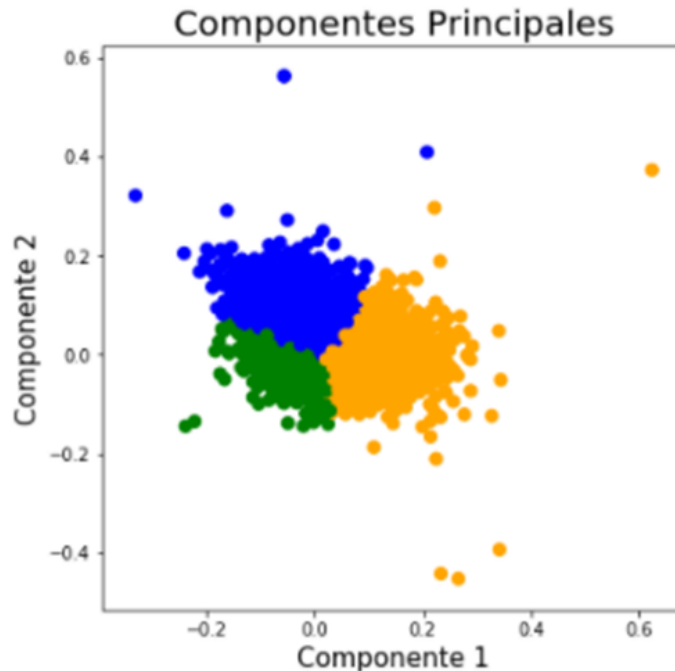


Figura 12. Componentes principales Python.

En R, se siguió un procedimiento análogo, utilizando las mismas variables ('table' y 'profundidad'). El código de Jambú indicó también tres clusters óptimos (Figura 13). Sin embargo, la visualización obtenida mediante PCA (Figura 14) mostró una menor definición entre los grupos, lo cual puede atribuirse a diferencias en la configuración del algoritmo o en el manejo de la normalización de datos entre ambos lenguajes.

```

> wss <- (nrow(df)-1)*sum(apply(df,2,var))
> for (i in 1:11) wss[i] <- sum(kmeans(df,
+   centers=i)$withinss)
> plot(1:11, wss, type="b", xlab="Numero de clusters",
+   ylab="Wcss")
.
```

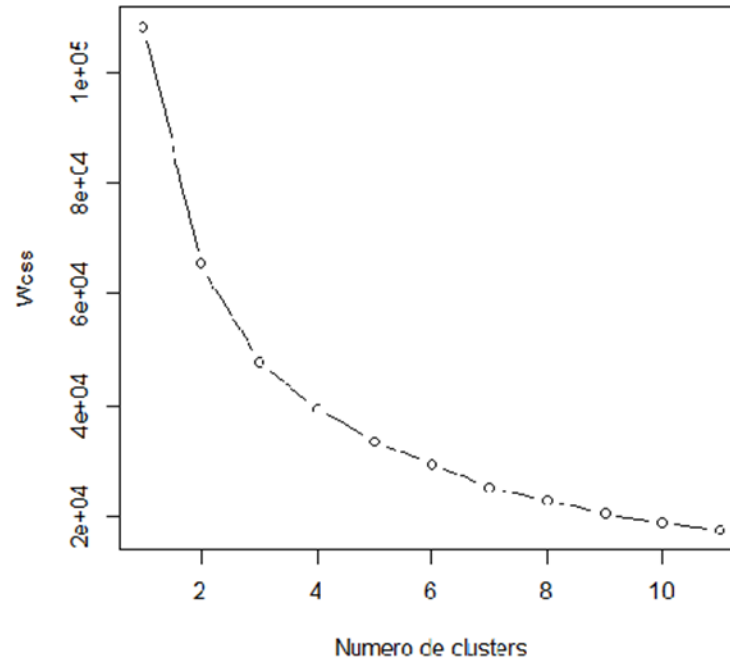
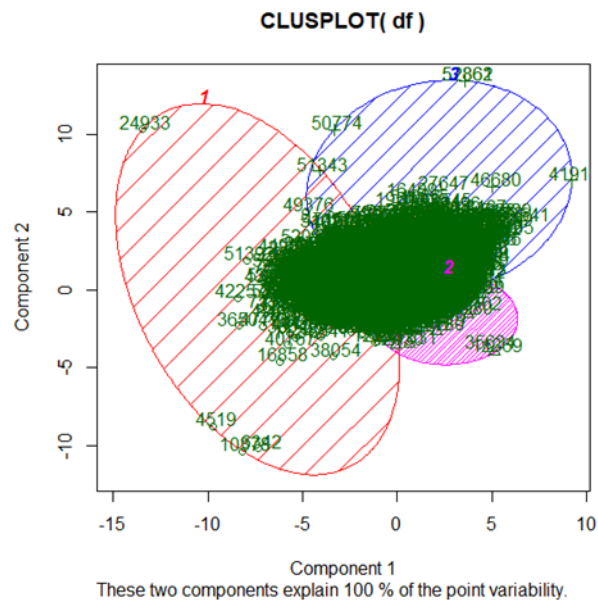


Figura 13. Codo de Jambú R.



These two components explain 100 % of the point variability.

Figura 14. Componentes principales R.

Para sintetizar los resultados obtenidos, en la Tabla 1 se presenta una comparación general entre los métodos aplicados en R y Python, considerando el nivel de precisión, la facilidad de uso y la calidad de las salidas gráficas.

Tabla 2. Comparación general entre R y Python.

Técnica	Software	Precisión (%)	Ventajas	Observaciones
k-NN	Python	82	Implementación flexible, buena visualización	Requiere ajuste manual de k
k-NN	R	83	Procesamiento simple y rápido	Menor flexibilidad en gráficos
Árboles de decisión	Python	84	Visualización clara, interpretación intuitiva	Configuración más extensa
Árboles de decisión	R	83	Fácil modelado y representación visual	Menor control en profundidad del árbol
Bosques aleatorios	Python	84	Alta precisión y estabilidad del modelo	Requiere mayor capacidad computacional
Bosques aleatorios	R	84	Resultados comparables con menor complejidad	Visualización menos detallada
Cluster k-means	Python	-	Gráficos definidos y separación clara de grupos	Alta consistencia en clustering
Cluster k-means	R	-	Resultados adecuados, pero menos diferenciados	Sensibilidad a la normalización

En términos generales, los resultados evidencian que tanto R como Python constituyen herramientas eficaces para la minería de datos. Python destaca por su versatilidad, capacidad de integración y calidad visual en la representación de resultados, mientras que R ofrece una mayor sencillez en la ejecución de modelos y una curva de aprendizaje más corta para usuarios con formación estadística. Ambos entornos complementan sus fortalezas y demuestran su valor en la enseñanza y la práctica aplicada del análisis de datos.

CONCLUSION

El análisis comparativo entre los entornos de programación R y Python permitió constatar que ambos constituyen herramientas robustas y complementarias para la minería de datos aplicada a la clasificación y segmentación de información. En los distintos algoritmos implementados —vecinos más cercanos,

árboles de decisión, bosques aleatorios y análisis de clusters—, los resultados mostraron niveles de precisión cercanos, con diferencias mínimas atribuibles principalmente a las particularidades de la sintaxis, la estandarización de datos y la gestión interna de los modelos.

Python destacó por su versatilidad, su integración con múltiples librerías de aprendizaje automático y su amplia capacidad para la visualización de resultados mediante entornos interactivos. En contraste, R evidenció una mayor accesibilidad y eficiencia en la ejecución de procedimientos estadísticos, ofreciendo una curva de aprendizaje más corta para los usuarios provenientes de disciplinas analíticas o académicas. En ambos casos, el carácter de software libre y la disponibilidad de documentación abierta constituyen un valor añadido que democratiza el acceso a metodologías avanzadas de análisis.

Los resultados confirman que, para tareas de clasificación supervisada, los algoritmos implementados en ambos lenguajes alcanzaron precisiones superiores al 80%, lo que valida su aplicabilidad en contextos reales de minería de datos. En los procedimientos de agrupamiento no supervisado, ambos entornos lograron identificar estructuras consistentes en los datos, aunque Python mostró una delimitación más clara de los grupos formados.

De manera general, este estudio evidencia que el uso combinado de R y Python amplía el espectro analítico disponible para el investigador, permitiendo seleccionar la herramienta más adecuada según los objetivos del estudio, el volumen de información o la naturaleza de las variables. Así, más que optar por uno u otro entorno, se propone un enfoque complementario que fomente la integración de ambos lenguajes como parte del perfil profesional del analista de datos contemporáneo.

REFERENCIAS

- Aguado, L. (2018). El análisis de datos en México y su aplicación en estrategias de mercado. *Revista Mexicana de Negocios*.
- Ballesteros, E., Iñiguez, R., & Velasco, M. (2018). *Fundamentos de la minería de datos*. Madrid: Universidad Autónoma.
- Berlanga, V., Rubio, M., & Vilà, R. (2013). *Árboles de decisión: aplicaciones prácticas*. Barcelona: UOC.
- Carvajal, M., & Maita, P. (2008). *Aplicaciones estadísticas con SPSS y Excel*. Universidad de los Andes.
- Figueras, M. (2001). *Métodos de clasificación y clustering*. Barcelona: Universitat Politècnica de Catalunya.
- Gatto, F. (2018). *INTRODUCCIÓN al lenguaje R y sus aplicaciones estadísticas*. Buenos Aires: Ediciones Científicas.

- Han, J., Pei, J., & Kamber, M. (2011). *Data Mining: Concepts and Techniques*. Morgan Kaufmann.
- International Business Machines. (2008). *Statistical Analysis Handbook*. IBM Corporation.
- Lenton, F. (2013). Análisis de votaciones nominales mediante R y Python. *Revista de Análisis Político*, 9(2), 45–67.
- Medina, L., & Ñique, A. (2017). Bosques aleatorios y su aplicación en la minería de datos. *Revista de Computación Aplicada*, 12(3), 22–35.
- Montoro, J. (2012). *Programación en Python: Fundamentos y aplicaciones*. Madrid: Anaya Multimedia.
- Moreno, J. (2018). El crecimiento exponencial de los datos digitales. *El Economista*.
- Python Software Foundation. (2020). *Python documentation*. Retrieved from

ÍNDICE NACIONAL DE PRECARIEDAD LABORAL: TRATAMIENTO ESTADÍSTICO DE BASES DE DATOS (ENOE)

Myrna Enedelia González Meneses, José Luis Hernández González, Mario Eduardo Leal López

Tecnológico Nacional de México/ITApizaco

myrnagm@yahoo.com, luis.hg@itapizaco.tecnm.mx, mario.ll@itapizaco.tecnm.mx

RESUMEN

Aunque han mejorado los algoritmos en la minería de datos, aún faltan esquemas adecuados para el tratamiento de información de carácter socio-económico; tales técnicas no son lo suficientemente avanzadas para discriminar información valiosa de las bases de datos existentes que han generado organismos gubernamentales o asociaciones privadas interesadas en la construcción de las políticas públicas. Para poder proporcionar un criterio adecuado de tratamiento de datos, se requiere establecer una metodología que permita discriminar y seleccionar variables para un óptimo tratamiento y codificación. Se presenta el marco teórico y metodología utilizada para analizar la información proveniente de la **Encuesta Nacional de Ocupación y Empleo (ENOE)** de los años 2005 al 2009, lo que permitió construir un Índice Nacional de Precariedad Laboral.

Palabras clave: SPSS, Precariedad Laboral, ENOE

INTRODUCCIÓN

Una de las problemáticas actuales más recurrentes en la gobernanza de cualquier país, se refiere principalmente a las condiciones precarias en que se encuentra el tema del trabajo; sin lugar a dudas, esta problemática es producto de la implementación de políticas públicas en el mundo globalizado. Las actuales formas de comercialización entre países, generan diferentes formas de trabajo, ya sea al interior de un país determinado o en alguna región en particular, por ello, diversos autores han propuesto definiciones de las condiciones laborales, las cuales, no han podido ser unificadas del todo; además, tampoco se cuenta con una metodología e información *exprofesa* para poder medir esas condiciones laborales.

Aunque actualmente se han desarrollado nuevas técnicas para el tratamiento de grandes cantidades de información, -algunas de ellas basadas en técnicas inteligentes- es muy complicado establecer un esquema de minería de datos adecuado y particularizado a cada situación, considerando el alto costo de generación

de información en problemáticas socio-económicas que requieren los gobiernos para establecer y mejorar sus políticas públicas.

Para desarrollar el procedimiento de cálculo del Índice de Nacional de Precariedad Laboral, se optó por utilizar la información generada por el Instituto Nacional de Estadística, Geografía e Informática (INEGI), particularmente de las Bases de Datos de la Encuesta Nacional de Ocupación y Empleo (ENOE); así como también, se decidió realizar el tratamiento de los datos, con algunas técnicas de estadística multivariada, con el software estadístico SPSS (Social Package for Social Sciences), con la finalidad de establecer un indicador para las condiciones laborales en nuestro país; para ello, fue de gran ayuda plantear y delimitar el marco teórico que sirviera como soporte para el manejo de la información proveniente de las bases de datos de la ENOE para los años 2005 al 2009.

MATERIALES Y MÉTODOS

Aunque la mayor parte de la población (ya sea que ésta cuente con algún nivel educativo o ninguno) no puede definir con certeza el concepto de trabajo precario, de acuerdo con las estadísticas y la definición de bienestar propuesta por el Consejo Nacional para la Evaluación de la Pobreza (CONEVAL): el bienestar económico, las carencias por falta de acceso a la salud, las carencias por falta de acceso a la seguridad social, entre otros factores, la precariedad laboral la vive día con día, una altísima cantidad de personas a nivel nacional. En la tesis de doctorado en Desarrollo Regional denominada: LA PRECARIEDAD LABORAL URBANA EN MÉXICO: UNA PERSPECTIVA REGIONAL-ESPACIAL 2005-2009 (González, 2011), se estableció la siguiente secuencia de actividades para la construcción de un Índice de Precariedad Laboral a nivel nacional:

- Planteamiento teórico-metodológico y planteamiento de la investigación.
- Trabajo decente, antecedentes y situación de la Seguridad Social en México.
- Construcción del Nacional Índice de Precariedad Laboral con las técnicas estadísticas multivariadas: análisis por componentes principales y conglomerados.
- Análisis espacial de la Precariedad Laboral en México con el Índice de Morán.
- Conclusiones y recomendaciones para la elaboración de políticas públicas.

De acuerdo con Moreno y estadísticas de la compañía Crowd Flower (2017), en una encuesta aplicada a 179 científicos de datos seleccionados en todo el mundo, se identificó que la distribución de las actividades que les toma mayor tiempo en su quehacer para el tratamiento de datos se encuentra repartida en los siguientes porcentajes:

- 51% coleccionar, etiquetar, limpiar y organizar los datos.
- 19% construir y modelar los datos.
- 10% modelación de datos para patrones.
- 9% refinar algoritmos.
- 8% otras actividades.

En dicha encuesta, se encontró que entre las actividades que los científicos más disfrutaban, están: construir y modelar los datos, aplicar minería de datos para encontrar patrones y el refinar algoritmos. Entre las menos gustadas están: limpiar y organizar datos, etiquetarlos y coleccionarlos. El 51% de los encuestados reportó que trabajan con datos no estructurados. Los datos con los que trabajan provienen principalmente de los sistemas internos de las compañías en las que trabajan, seguido de los que coleccionan de forma manual y, por último, de los conjuntos de datos disponibles públicamente.

Son tres áreas en las que principalmente se desarrollan los científicos de datos:

- ☐ *Big Data* para procesar datos,
- ☐ Minería de datos para analizar e identificar relaciones ocultas, patrones y tendencias,
- ☐ Visualización de datos para explicar y socializar mejor la información obtenida.

La información anterior es importante, en virtud de que la ciencia de datos, requiere de conocimientos, habilidades, experiencia y software especializado para lograr obtener el mayor provecho en la extracción, visualización y procesamiento de grandes cantidades de datos, así como también, solventar las limitantes y problemáticas que estos procesos conllevan.

Aunado con la cantidad de recursos tanto económicos como humanos requeridos para realizar la obtención y recolección de datos, se determinó una mayor relevancia en establecer un marco teórico adecuado que permitiera construir una definición operativa del trabajo precario, acorde al producto laboral de la población económicamente activa.

Trabajo atípico: Son aquellas modalidades de contratación laboral que no se insertan dentro del modelo de relación laboral normal, es decir, son aquellas que no se encuentran reguladas jurídicamente o en donde no se cumplen todas las características de la relación laboral normal (Caamaño, 2008).

Trabajo precario: Se define en términos del cumplimiento o no de la regulación del Estado sobre las condiciones de la venta y uso de la fuerza de trabajo (Rojas y Salas, 2008).

De aquí se desprenden los conceptos básicos que sirvieron para discriminar las variables de las bases de datos de la ENOE; en los siguientes párrafos, se describe la constitución de la información que proporcionan los cuestionarios del sector laboral a nivel nacional.

- 1) *El grado de certidumbre respecto de la continuidad de un trabajo*, los trabajos precarios son de corta duración respecto de los cuales, el riesgo de despido es elevado. Esta dimensión engloba igualmente, todo trabajo irregular en la medida en que existe incertidumbre respecto de su renovación;
- 2) *El control sobre las condiciones de trabajo*, cuanto menos controle el trabajador -ya sea individual o colectivamente- sus condiciones de trabajo, de salario o de ritmo de trabajo, menos asegurada está la continuidad del empleo;
- 3) *La protección social*, este es un aspecto crucial, es decir, saber en qué medida los trabajadores se encuentran protegidos, ya sea por la ley o por alguna organización colectiva;
- 4) *Un ingreso estable*, Rodgers (1989), reconoce como ambiguo este concepto relativo a los ingresos, ya que los empleos con remuneraciones bajas pueden ser considerados precarios si se encuentran asociados a un contexto de pobreza o de inserción social desfavorable.

Asimismo, se definieron tres objetivos clave con la finalidad de obtener una metodología que pudiera ser aplicable a cada base de datos, la cual, se aplica a través de cuestionarios puerta a puerta de hogares seleccionados y corresponde a la aplicación de manera trimestral cada año, de aquí que se generan un sinnúmero de variables y esquemas de respuestas que se capturan en archivos de bases de datos (González, 2011).

- Operativizar la definición de precariedad laboral bajo las condiciones actuales de trabajo a nivel nacional.
- Construir un modelo estadístico multivariado aplicando el Análisis por Componentes Principales, el cual permite manipular factores socio-demográficos y económicos, para evaluar la precariedad laboral a nivel estatal mediante un Índice de Precariedad Laboral nacional en el período comprendido del año 2005 al año 2009.
- Realizar un análisis regional-espacial geo-referenciado de la precariedad laboral y el crecimiento económico de los años 2005 al 2009.

Estadísticas laborales para medir la ocupación en México

A continuación, se muestra la manera en que se construyeron las bases de datos. A partir del año 1985, ha mejorado considerablemente la cobertura geográfica y espacial de los censos económicos del Instituto Nacional de Estadística, Geografía e Informática (INEGI) en los sectores de manufactura, comercio, pesca, servicios financieros. Igualmente, es el caso de las encuestas. Para este estudio, se seleccionó primordialmente, la información de las Bases de Datos de la Encuesta Nacional de Ocupación y Empleo (ENOE) que inició a partir del año 2005.

ANTECEDENTES

La *Encuesta Nacional de Ocupación y Empleo (ENOE)* es la consolidación y fusión de la *Encuesta Nacional de Empleo Urbano (ENEU)* y la *Encuesta Nacional de Empleo (ENE)*, que por más de 20 años, estuvieron proporcionando información sobre la población ocupada y desocupada en nuestro país.

A su vez, éstas son resultado de una larga tradición de encuestas en hogares que dio inicio desde el año 1972 con el levantamiento de la *Encuesta Nacional de Hogares (ENH)* continuando en 1973; durante 1973 y 1974 la *Encuesta Continua de Mano de Obra (ECMO)*; de 1974 a 1984 la *Encuesta Continua sobre Ocupación (ECSO)*; de 1983 a 2004 la *Encuesta Nacional de Empleo Urbano (ENEU)*, de 1991 a 2004 la *Encuesta Nacional de Empleo (ENE)* y de 2005 a 2009 la *Encuesta Nacional de Ocupación y Empleo (ENOE)*

La ENOE, es resultado de una evaluación integral de varios años, durante los cuales el Instituto Nacional de Estadística y Geografía (INEGI), ha realizado una serie de actividades de carácter conceptual, metodológico y de mejora de procesos, con la finalidad de presentar una nueva encuesta que permite captar y conocer de mejor manera las características del mercado laboral mexicano. Entre dichas actividades, se encuentra la realización de talleres de consulta con usuarios tanto del sector académico como de instituciones públicas y un foro con personal de la Organización para la Cooperación y el Desarrollo Económico (OCDE); así como la ejecución de diferentes pruebas para los instrumentos de captación y los procesos de trabajo, entre ellas, una de significancia estadística con la idea de confrontar los resultados obtenidos con una análoga en dimensión y características extraída del levantamiento rutinario ENE-ENEU.

Los cuestionarios de la ENOE (Sociodemográfico y de Ocupación y Empleo) dan respuesta a los cambios que ha experimentado el panorama laboral en México, se sustentan en un nuevo marco conceptual de referencia, lo que permite, una caracterización de los mercados de trabajo en México así como del trabajo independiente y de la multiocupación, al tiempo que brinda más elementos para analizar el trabajo subordinado y remunerado, el contexto bajo el cual los individuos se incorporan a un empleo y también bajo qué circunstancias lo pierden. Un punto no menos esencial en el nuevo diseño, es su capacidad para captar la búsqueda de empleo y cuantificar mejor la presión ejercida en el mercado laboral tanto por parte de los desocupados como por parte de quienes ya cuentan con una ocupación pero que buscan otra y compiten con los primeros por una plaza de trabajo. Los cambios de diseño permiten, asimismo, calcular tasas de desocupación bajo los estándares establecidos por OCDE y profundizar en fenómenos tales como los que atañen a grupos de población quienes han desistido de buscar un empleo o a otros más para los

cuales hay un conflicto entre sus tareas con respecto al hogar por un lado y la necesidad de trabajar por el otro.

Con respecto a los conceptos básicos de la ENOE, su organización operativa, el diseño conceptual y estadístico, el levantamiento de la información, la filosofía, conceptos y fases en el tratamiento de la información, la comunicación operativa y concertación, la conformación, integración y preparación de las bases de datos, la difusión de resultados, la periodicidad de las encuestas, entre otros aspectos importantes para conocer más a detalle el funcionamiento de esta encuesta.

Contenidos básicos, Métodos, Procedimientos y Contenido de las Bases de Datos de la ENOE

Las encuestas de la ENOE son continuas a nivel nacional, y a las personas residentes habituales de las viviendas seleccionadas, se les entrevista una vez por trimestre durante cinco de éstos. El diseño conceptual de la ENOE, incluye los siguientes indicadores:

POBLACIÓN OBJETIVO

- Las personas residentes habituales de las viviendas seleccionadas.
- Para las características económicas, se captan datos de la población de 12 y más años de edad; aunque los indicadores que se difunden mensual y trimestralmente, se generan para la población de 14 y más años de edad.

COBERTURA TEMÁTICA

- Características socio-demográficas de los miembros del hogar.
 - Condición de residencia.
 - Parentesco.
 - Sexo y edad.
 - Lugar de nacimiento.
 - Educativas.
 - Número de hijos (sólo para mujeres de 12 años y más de edad).
 - Estado conyugal.
- Población económicamente activa (PEA)
 - Población ocupada
 - Ocupación.
 - Condición de multiocupación.
 - Datos de trabajo secundario (ocupación, sector de actividad, acceso a la atención médica).

- Posición en la ocupación.
 - Tamaño del establecimiento.
 - Tipo de contrato.
 - Afiliación sindical.
 - Sector de propiedad.
 - Sector de actividad.
 - Disponibilidad de local.
 - Tipo de local.
 - Horas trabajadas.
 - Remuneraciones al trabajo, formas y periodo de pago.
 - Prestaciones laborales.
 - Acceso a servicio médico.
 - Datos sobre trabajo secundario.
 - Presión laboral.
 - Regularidad en el trabajo principal.
- Población desocupada
 - Tipo de trabajo buscado.
 - Duración de la desocupación.
 - Experiencia laboral.
 - Razones de la desocupación.
 - Posición en la última ocupación.
 - Ocupación en el último trabajo.
 - Sector de actividad en el último trabajo.
- Población no económicamente activa (PNEA)
 - Razones de no actividad.
 - Motivos de desaliento.
 - Experiencia laboral.
 - Razones de abandono del último trabajo.
 - Posición en la última ocupación.
 - Ocupación.
 - Sector de actividad.

CLASIFICADORES UTILIZADOS

- Catálogo e Instructivo de Codificación de Parentesco, 2008.
- Catálogo e Instructivo de Codificación de Lugar de Nacimiento, 2005.
- Catálogo de Codificación de Carreras, 2005.
- Catálogo de Dependencias e Instituciones de Interés Público (CADIIP), 2007.
- Catálogo de Franquicias en México 2006.
- Clasificación Mexicana de Ocupaciones (CMO), volumen I y II, 2005.
- Instructivo para la codificación de carreras.
- Sistema de Clasificación Industrial de América del Norte (SCIAN), versión hogares, 2007.

INSTRUMENTOS DE CAPTACIÓN

- Cuestionario sociodemográfico.
- Cuestionario de ocupación y empleo, básico y ampliado.

Corte temporal: Mensual y trimestral.

Cobertura geográfica: La encuesta permite referenciar cortes geográficos de la información en los niveles:

- Nacional.
- Entidad federativa.
- 32 ciudades autorrepresentadas.
- Localidades de 100 000 y más habitantes.
- Localidades de 15 000 a 99 999 habitantes.
- Localidades de 2 500 a 14 999 habitantes.
- Localidades con menos de 2 500 habitantes.

Diseño estadístico

Tamaño de la muestra

Está calculado desde el punto de vista del indicador de mayor exigencia en términos de tamaño muestral. Lo anterior garantiza que las estimaciones del resto de las variables de interés queden cubiertas con este tamaño. Trimestralmente, el tamaño de la muestra es de aproximadamente 120,260 viviendas.

Unidad de observación: El hogar

Unidad de muestreo: La vivienda particular

Unidad de análisis: El hogar y los residentes de la vivienda.

Marco muestral

El marco muestral utilizado para la ENOE es el marco nacional de viviendas del año 2002 del INEGI, construido a partir de la información cartográfica y demográfica obtenida durante el XII Censo General de Población y Vivienda, 2000.

Esquema de muestreo

El tipo de muestreo utilizado es probabilístico, bietápico, estratificado y por conglomerados.

Periodo de levantamiento

Semanalmente de lunes a domingo, durante todo el año.

El cuestionario socio-demográfico ampliado, incluye once aspectos y/o baterías relevantes:

Las preguntas que incluye el cuestionario socio-demográfico y de ocupación y empleo ampliado de la ENOE, se encuentran en las siguientes once baterías (Tabla 1):

Tabla 1. Clasificación y Número de Batería (por sección) del cuestionario de la ENOE

BATERÍA (por sección)	No. preguntas
I. CONDICIÓN DE OCUPACIÓN	6
II. NO OCUPADOS	9
III. CONTEXTO LABORAL	21
IV. CARACTERÍSTICAS DE LA UNIDAD ECONÓMICA	10
V. JORNADA Y REGULARIDAD LABORAL	9
VI. INGRESOS Y ATENCIÓN MÉDICA	5
VII. TRABAJO SECUNDARIO	5
VIII. BÚSQUEDA DE OTRO TRABAJO	3
IX. ANTECEDENTES LABORALES	15
X. APOYOS ECONÓMICOS	2
XI. OTRAS ACTIVIDADES	1
total	86

RESULTADOS

Tabla 2. Índice de Precariedad Laboral de los años 2005 al 2009

Estado	IPL 2005	IPL 2006	IPL 2007	IPL 2008	IPL 2009
Aguascalientes	1.34	1.37	1.19	1.18	1.18
Baja California	1.26	1.09	0.88	0.95	0.96
Baja California Sur	1.27	1.19	1.07	1.15	1.15
Campeche	1.59	1.71	1.45	1.46	1.50
Coahuila	1.15	1.13	0.88	0.95	0.97
Colima	1.54	1.60	1.47	1.42	1.45
Chiapas	2.01	2.33	2.01	1.99	2.05
Chihuahua	1.17	1.00	0.84	0.91	0.91
Distrito Federal	1.37	1.42	1.16	1.22	1.23
Durango	1.41	1.39	1.20	1.13	1.16
Guanajuato	1.64	1.76	1.56	1.62	1.62
Guerrero	1.66	1.88	1.75	1.78	1.81
Hidalgo	1.84	2.08	1.94	2.01	2.05
Jalisco	1.46	1.53	1.31	1.37	1.37
México	1.48	1.69	1.48	1.43	1.45
Michoacán	1.80	2.09	1.94	2.01	2.06
Morelos	1.62	1.85	1.72	1.81	1.79
Nayarit	1.62	1.80	1.61	1.62	1.65
Nuevo León	1.19	1.08	0.92	0.96	0.96
Oaxaca	1.78	2.02	1.88	1.96	1.99
Puebla	1.84	2.18	2.02	2.07	2.11
Querétaro	1.40	1.41	1.17	1.22	1.23
Quintana Roo	1.34	1.43	1.19	1.27	1.28
San Luis Potosí	1.48	1.68	1.50	1.61	1.62
Sinaloa	1.52	1.68	1.44	1.47	1.50
Sonora	1.33	1.21	1.03	1.18	1.20
Tabasco	1.68	1.80	1.68	1.61	1.64
Tamaulipas	1.20	1.17	1.17	1.14	1.16
Tlaxcala	1.76	2.13	1.98	2.00	2.04
Veracruz	1.66	2.01	1.77	1.78	1.82
Yucatán	1.67	1.65	1.47	1.48	1.54
Zacatecas	1.63	1.90	1.73	1.76	1.80

Elaboración propia a partir de las Bases de Datos de la ENOE con componentes principales

CONCLUSIONES

Los resultados que arroja el Índice Nacional de Precariedad Laboral (Tabla 2) se pueden interpretar de la siguiente manera:

Los únicos Estados donde se visualiza una disminución constante del Índice de precariedad laboral durante el período 2005 al 2009 son Aguascalientes y Nuevo León; los Estados con un valor cercano a 2.00 son: Chiapas, Hidalgo, Michoacán, Oaxaca, Puebla y Tlaxcala, que son a los que consideramos con mayor precariedad y en menor medida Veracruz. El mínimo valor se encuentra en el Estado de Chihuahua para el año de 2007 y el máximo valor es de 2.33 para Chiapas en el año 2006. Se determinó la desviación estándar siendo la mínima para Tamaulipas de 0.0217 y el máximo lo presenta Baja California con 0.1502. Los Estados que manifestaron un promedio mayor a 2.00 en promedio para los cinco años fueron Chiapas y Puebla.

Los Estados que presentan un crecimiento del valor de precariedad a partir de 2007 son: Baja California, Coahuila, Campeche, Chihuahua, Distrito Federal, Guanajuato, Guerrero, Hidalgo, Jalisco, Michoacán, Nayarit, Oaxaca, Puebla, Querétaro, Quintana Roo, San Luis Potosí, Tlaxcala, Veracruz y Yucatán.

El mínimo valor de precariedad para el 2005, fue para Coahuila de 1.15 y el máximo para Puebla con 1.84. Para el año de 2009, el mínimo fue para Chihuahua con 0.91 y el máximo de 2.11 para Puebla.

En la actualidad, existe una gran cantidad de técnicas y estrategias para la generación de conjuntos de datos; es decir, se cuenta con diferentes dispositivos que a nivel industrial, nos permiten controlar un sinnúmero de variables: temperatura, presión, intensidad de luz, entre otras, y la medición se puede hacer en incrementos de tiempo muy pequeños como los acelerogramas de los sismos. Por otro lado, también se cuenta con la información que producen algunos dispositivos de consulta como todos aquellos que permiten recolectar datos de la Web y/o de software específico; por ejemplo el número de veces que se hace click en ciertas plataformas.

Sin embargo, existen diferentes tipos y fuentes de información de carácter más completo y que provienen de la consulta de información social relacionada con problemáticas de diversa índole, donde es de suma importancia conocer y establecer metodologías para proponer en la elaboración de planes de desarrollo en los países. Asimismo, como las cantidades de datos cada día se tornan más complejas y extensas, se ha hecho necesario tratar esa información con las técnicas actuales de la inteligencia artificial. Es importante resaltar, que además del tiempo que se consume en la recolección de datos, ahora se requiere destinar tiempo en establecer estrategias del tratamiento de datos, así como manipulación de variables y archivos, ya que, aunque actualmente se cuenta con equipos de cómputo sofisticados con una gran capacidad de

memoria y almacenamiento, las bases siguen creciendo y se deben establecer estrategias que nos permitan manipularlas adecuadamente y en forma más rápida y simple.

REFERENCIAS

- Caamaño Rojo, Eduardo (2008). Las transformaciones del trabajo, la crisis de la relación laboral normal y el desarrollo del empleo atípico. <http://blog.pucp.edu.pe/item/26493>
- Encuesta Nacional de Ocupación y Empleo (2005). 50 Preguntas y Respuestas. INEGI. Aguascalientes, Ags.
- Encuesta Nacional de Ocupación y Empleo (2005). Encuesta Nacional de Ocupación y Empleo. ENOE: Una nueva Encuesta para México. INEGI. Aguascalientes, Ags.
- Encuesta Nacional de Ocupación y Empleo (2007). ¿Cómo se hace la ENOE? Métodos y Procedimientos. INEGI. Aguascalientes, Ags.
- Encuesta Nacional de Ocupación y Empleo (2007). Conociendo la Base de Datos de ENOE. INEGI. Aguascalientes, Ags.
- González Meneses, Myrna Enedelia (2011). La Precariedad Laboral Urbana en México: una Perspectiva Regional-Espacial 2005-2009, Tesis de Grado, Doctorado en Desarrollo Regional. San Pablo Apetatitlán, Tlaxcala.
- Moreno Salinas, José (2017). Científico de Datos: Codificando el Valor Oculto e Intangible de los Datos. Revista Digital Universitaria. Vol. 18, Núm. 7, tomado de: <http://revista.unam.mx/vol.18/num7/art53/index.html>, el 22 de febrero de 2018

SERVICIOS ECOSISTÉMICOS EN EL SECTOR DE USO DE SUELO: ¿EXISTE EVIDENCIA DE SU CONTRIBUCIÓN AL BIENESTAR?

Martín López Ramírez, Juan Ruiz Ramírez

martinalr7@gmail.com, jruiz@uv.mx

RESUMEN

El sector de uso de suelo (AFOLU) se considera clave para impulsar las metas de desarrollo sostenible. Sin embargo, el sector afronta el reto de incrementar significativamente su producción con el fin de alimentar una población mundial creciente sin degradar más la base de los recursos naturales sobre los que los sistemas de producción dependen. En este sentido, opciones ecológicamente más viables han sido propuestas como medidas integrales para afrontar los retos de seguridad alimentaria y sostenibilidad del sector. No obstante lo anterior, la contribución de los ecosistemas forestales a la producción de alimentos y a los medios de subsistencia en general es aún pobremente entendida. Con el fin de contribuir estrechar esta brecha de conocimiento, este trabajo de investigación analizó la contribución de los ecosistemas forestales a la productividad del sector de uso de suelo en la región tropical. Basados en el estado de conservación de los ecosistemas forestales, se analizó el grado al cual la conservación del capital natural explica el comportamiento del PIB agrícola en el periodo 1990-2015. Lo anterior permitió inferir acerca de la contribución del capital natural a la productividad del sector y la importancia de los servicios ecosistémicos derivados de los bosques. El análisis reveló que los ecosistemas forestales juegan un papel fundamental en la creación de bienestar de manera directa e indirecta. No obstante, también se observó que existen incentivos privados a la deforestación. Ante la contraposición de intereses colectivos e individuales, la regulación del sector en la región estudiada resulta imperativa.

INTRODUCCIÓN

El sector de uso de suelo (AFOLU, por sus siglas en inglés) se considera clave en términos de seguridad alimentaria, política ambiental en el contexto de cambio climático y, en última instancia, como un elemento estratégico para impulsar el desarrollo sustentable (Di Gregorio *et al.*, 2017; Smith *et al.*, 2014). Desde la perspectiva de la seguridad alimentaria, el sector afronta el reto de incrementar significativamente su producción con el fin de alimentar una población mundial creciente (Godfray *et al.*, 2010). Se estima que para el año 2050 la producción de alimentos deberá incrementarse más del 60% con

respecto a sus niveles actuales para alimentar una población proyectada de 9.3 mil millones de habitantes. Sin embargo, dicho incremento debe lograrse sin degradar más la base de los recursos naturales sobre los que los sistemas de producción dependen (FAO, 2014). Lo anterior, supone una transformación integral de los sistemas productivos, ya que una estrategia basada únicamente en la expansión de la frontera agrícola en detrimento de los ecosistemas forestales pondría en entredicho la sostenibilidad del sector.

En este sentido, opciones ecológicamente más viables, como lo son la intensificación agrícola y el uso de tierras no forestales para fines productivos, han sido propuestas como medidas integrales para afrontar los retos de seguridad alimentaria y sostenibilidad del sector (Carter *et al.*, 2015). Sin embargo, la contribución de los ecosistemas forestales a la producción de alimentos y a los medios de subsistencia en general es aún pobremente entendida (Reed *et al.*, 2017). Con el fin de contribuir estrechar esta brecha de conocimiento, este trabajo de investigación analizó la contribución de los ecosistemas forestales a la productividad del sector de uso de suelo, haciendo énfasis en la región tropical.

En concreto, basados en el estado de conservación de los ecosistemas forestales (aproximado mediante la etapa de transición forestal en que los países se han clasificado durante el periodo 1990-2010, ver Hosonuma *et al.*, (2012)), se analizó mediante técnicas estadísticas, el grado al cual la conservación del capital natural mencionado explica el comportamiento del PIB agrícola en los países estudiados en el periodo 1990-2015. Lo anterior permitió inferir acerca de la contribución del capital natural a la productividad del sector y la importancia de los servicios ecosistémicos derivados de los bosques.

Servicios Ecosistémicos y capital natural

Los servicios ecosistémicos se definen como los beneficios que los seres humanos obtienen de ecosistemas funcionales. Dichos beneficios se derivan de procesos y funciones ecológicas y pueden contribuir al bienestar humano de manera directa o indirecta, consciente o inconsciente (Costanza *et al.*, 2017). Los servicios ecosistémicos se clasifican en servicios de provisión o bienes ecosistémicos, servicios de soporte, servicios de regulación y servicios culturales (MA, 2003). Un ejemplo de la primera categoría es la producción de alimentos y maderas. Un ejemplo de la segunda categoría son la fotosíntesis y la producción primaria neta. Ejemplos de la tercera categoría son la regulación de los ciclos hídricos, regulación del microclima y la formación y conservación de suelo. Ejemplos de la cuarta categoría son la recreación en áreas naturales y la conservación de animales y/o plantas sagradas (Costanza *et al.*, 2017; Locatelli, Lavorel, Sloan, Tappeiner, & Geneletti, 2017; Raudsepp-Hearne *et al.*, 2010).

Los ecosistemas que proveen los servicios ecosistémicos son referidos como capital natural, bajo el entendido de que el capital es un stock que produce un flujo de servicios a través del tiempo (Costanza *et al.*, 2017). Para promover el bienestar humano, el capital natural debe interactuar con otras formas de capital; como lo son el capital físico (fabricado por el hombre), humano, social o cultural. Por lo tanto, los servicios ecosistémicos son la contribución relativa del capital natural al bienestar humano, tomando en cuenta su interacción con otras formas de capital (Costanza *et al.*, 2017).

Es importante señalar que el vínculo entre procesos y funciones ecológicas y el bienestar humano es poco entendido en la actualidad. Lo anterior se debe en gran parte a que la interacción entre el capital natural y otras formas de capital es crucial para que los procesos y funciones ecológicas alcancen la dimensión de bienestar humano. Es decir, que dichos procesos y funciones se consideren servicios ecosistémicos y no únicamente relaciones biofísicas. Sin embargo, las combinaciones entre diferentes tipos de capital que resultan en una provisión sustentable de servicios ecosistémicos son sumamente complejas (Costanza *et al.*, 2017; Raudsepp-Hearne *et al.*, 2010) y los métodos y herramientas para evaluar dichas relaciones son aún limitadas (IPCC, 2014).

MATERIALES Y MÉTODOS

Unidades observacionales

Las unidades observacionales son los países de la franja tropical clasificados dentro del modelo de transición forestal por Hosonuma *et al.* (2012). Dicho estudio consideró a los países no pertenecientes al anexo I de la Convención Marco de las Naciones Unidas sobre el Cambio Climático. En total el trabajo citado consideró 100 países localizados en América, África y Asia. Sin embargo, debido a la falta de información, en este estudio se consideraron 96 de los 100 países estudiados.

Variables y fuentes de información

Las variables utilizadas para ajustar el modelo de regresión, descrito posteriormente, son las siguientes:

- **PIB agrícola (PIB):** incluye la producción de los componentes de silvicultura, caza, pesca, cultivos y ganadería. Se obtiene al considerar el valor agregado de la producción de los componentes mencionados después de sustraer el valor de los insumos intermedios. El indicador se mide en dólares americanos a valores constantes de 2010 y se obtuvo de la base de datos del banco mundial (World Bank, 2018a).
- **Cobertura forestal (F):** considera la superficie con árboles de al menos cinco metro de altura *in situ* en bosques naturales o plantados. Excluye árboles en sistemas de producción agrícola. El

indicador se mide en kilómetros cuadrados y se obtuvo de la base de datos del banco mundial (World Bank, 2018c).

- **Cobertura forestal relativa (FR):** es la proporción de la cubierta forestal con respecto a la superficie total del país (se excluyen cuerpos de agua interiores de la superficie total) medida en términos porcentuales. Este indicador se obtuvo de la base de datos del banco mundial (World Bank, 2018c).
- **Superficie agrícola (A):** Se refiere a la superficie de tierra cultivable, cultivos perenes y en pastizales permanentes medida en kilómetros cuadrados. Este indicador se obtuvo de la base de datos del banco mundial (World Bank, 2019).
- **Fase (P):** corresponde a una de las cuatro fases (pre-transición, transición temprana, transición tardía, postransición) del modelo de transición forestal. Dicha variable se obtuvo de Hosonuma *et al.* (2012) y se determinó en el estudio original en base a la cobertura forestal y tasas de deforestación observadas en el periodo 1990-2010.
- **Continente (C):** América, África o Asia, dependiendo de la localización del país. Dicha variable se obtuvo de Hosonuma *et al.* (2012).

Cabe mencionar que por cada unidad de estudio (país) en el periodo analizado (1990-2015) se obtuvo un registro de cada una de las primeras cuatro variables. Las últimas dos variables tienen un registro determinado únicamente por unidad de estudio, es decir, no sufren modificaciones a través del tiempo.

Modelo de regresión lineal

Se partió de la hipótesis de que el PIB agrícola es una función del capital natural (la cual se aproxima mediante las variables F y P) y la superficie en producción (aproximada mediante A). En específico, se consideró una función de producción Cobb-Douglass expresada como:

$$Y = \alpha_0 A^{\alpha_1} F^{\alpha_2} \quad (3)$$

Donde:

Y es el PIB agrícola;

α_j son parámetros a estimarse.

En el caso de la ecuación (3), el parámetro α_0 se interpreta como el nivel tecnológico o escala de producción y los parámetros α_1 y α_2 como las elasticidades parciales de los respectivos insumos. La escala de producción indica la cantidad de producto que se obtiene al usar una unidad de cada insumo, mientras

que las elasticidades parciales indica el cambio porcentual en la producción debido a un cambio de uno por ciento en la variable correspondiente (Gujarati, Guerrero, & Medina, 2004; Varian, 2010).

Para poder realizar una estimación de los parámetros, se procedió a linealizar la ecuación (3) mediante la aplicación de logaritmos. Una mejor aproximación a la ecuación (3) considera variables adicionales y toma la siguiente forma:

$$\log Y_{it} = \beta_0 + \beta_1 \log A_{it} + \beta_2 \log F_{it} + \beta_3 \log A_{it} \log F_{it} + \beta_4 FR_{it} + \beta_5 P + \beta_6 C + u_{it} \quad (4)$$

Donde:

Los subíndices i y t denotan el país y el año respectivamente;

β_j son los coeficientes a estimarse;

u representa los residuales del modelo.

Como se desprende de la ecuación (4), los coeficientes β_5 y β_6 impactan el intercepto del modelo dependiendo de la fase y la ubicación del país correspondiente. Lo que equivale a decir que la escala de producción (parámetro α_0 de la ecuación (3)) se determinó por las distintas combinaciones de las variables P y C . De manera similar, el modelo estimado es log-lineal en las variables Y , A y F , por lo que los coeficientes β_1 , β_2 y β_3 determinan de manera directa las elasticidades parciales de las variables A y F . Como se mostrará más adelante, el coeficiente β_4 influye también la estimación de elasticidad parcial de F . Es decir que se asumió que las elasticidades parciales (parámetros α_2 y α_3 de la ecuación (3)) se determinan por la composición del uso de suelo, reflejado en las variables A , F y FR .

Para ajustar el modelo log-lineal presentado en la ecuación (2), se usó el método de mínimos cuadrados ordinarios, el cual se implementó en el software R 3.5.2.

RESULTADOS

Ajuste del modelo regresión

Los coeficientes del modelo de regresión y sus propiedades estadísticas de interés se muestran en la Tabla 6. Como se puede apreciar, todos ellos resultaron altamente significativos ($\alpha = 0.01$). Aunado a lo anterior, se obtuvo una R^2 de 0.85, lo cual demuestra que el modelo ajustado explica una fracción importante de la varianza observada en los datos. Estos dos hechos en conjunto sugieren que se obtuvo un ajuste robusto a través del modelo de regresión seleccionado.

Como se mencionó con anterioridad, el objetivo del modelo de regresión fue estimar los parámetros de una función de producción Cobb-Douglass (ver ecuación (1)). En las siguientes secciones se muestra como los coeficientes del modelo de regresión fueron utilizados para estimar los parámetros de la función mencionada; es decir, la escala de producción y las elasticidades parciales de los distintos usos de suelo, así como sus implicaciones prácticas.

Tabla 6. Coeficientes del modelo de regresión

Variable	Coeficiente	Valor del coeficiente	Error estándar	IC al 95%
Intercepto	β_0	6.4895**	0.0994	6.29, 6.68
Log A	β_1	0.1759**	0.0279	0.12, 0.23
Log F	β_2	0.1536**	0.0342	0.09, 0.22
Log A Log F	β_3	0.0486**	0.0053	0.04, 0.06
FR	β_4	-0.0079**	0.0006	-0.009, -0.006
Fase: transición temprana	β_5	0.1409**	0.0253	0.09, 0.19
Fase: transición temprana	β_5	0.2593**	0.0314	0.19, 0.32
Fase: postransición	β_5	0.3470**	0.0335	0.28, 0.41
Continente: América	β_6	0.3874**	0.0186	0.35, 0.42
Continente: Asia	β_6	0.7017**	0.0204	0.66, 0.74

Fuente: elaboración propia.

*Nota: ** indica que el coeficiente es altamente significativo con $\alpha = 0.01$.*

Los niveles de referencia son África y la fase de pre transición

Elasticidades parciales

A partir del modelo estimado (ver ecuación (2)), la función de producción en su forma no lineal se expresa de la siguiente manera:

$$Y_i = \alpha_0(P, C)A^{\beta_1}F^{\beta_2}(AF)^{\beta_3}FR^{\beta_4}10^{u_i} \quad (5)$$

Donde:

Y_i es el PIB agrícola del i-esimo país¹;

$\alpha_0(P, C)$ es la escala de producción en función de las variables fase y continente (ver detalles en sección previa);

β_j son los coeficientes estimados (ver Tabla 6).

¹ Los índices de tiempo t fueron omitidos por simplicidad.

Como se puede observar, la elasticidad parcial de la superficie agrícola (parámetro α_1 de la ecuación 1) está determinado por los coeficientes β_1 y β_3 . En particular, la suma de dichos coeficientes determina la elasticidad parcial de la superficie agrícola. De modo similar, la elasticidad parcial de la cubierta forestal (parámetro α_2 de la ecuación 1) está determinada por los coeficientes β_2 , β_3 y β_4 . Para ver lo anterior se parte de la definición de cubierta forestal relativa, la cual se expresa matemáticamente como $\frac{F}{S}$ (World Bank, 2018b), donde S denota a la superficie total del país. Una simple manipulación de la ecuación (3) muestra que esta se puede expresar de la siguiente manera:

$$Y = \alpha_0(P, C)S^{-\beta_4}A^{\alpha_1}F^{\alpha_2} \quad (6)$$

Donde:

$\alpha_1 = \beta_1 + \beta_3$ es la elasticidad parcial de la superficie agrícola;

$\alpha_2 = \beta_2 + \beta_3 + \beta_4$ es la elasticidad parcial de la cubierta forestal.

La estimación numérica puntual y por intervalo de las elasticidades parciales, dadas las transformaciones antes expuestas, se muestran en la Figura 5. Las propiedades económicas de la función de producción saltan a la vista una vez obtenidas las estimaciones mencionadas: para comenzar, ambos usos de suelo muestran un producto marginal decreciente, ya que sus elasticidades parciales en cualquier caso son menores a la unidad. Esto indica que al incrementar el área de un uso de suelo en específico, el PIB responde de manera positiva pero a una tasa decreciente. Además, la función de producción estimada presenta rendimientos decrecientes a escala, lo cual se debe a que en todos los casos, la suma de las elasticidades parciales es menor a la unidad. Los rendimientos a escala representan el grado al que responde la producción al incrementar todos los factores en una misma escala. En este caso, el incremento simultáneo de la superficie agrícola y forestal por un escalar s , resulta en un incremento menor que proporcional (menor a s) en el PIB agrícola.

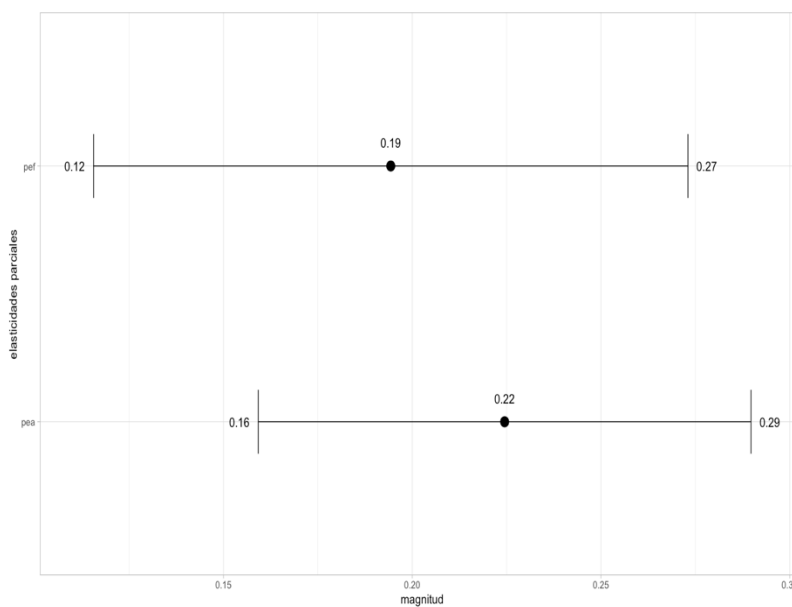


Figura 5. Estimaciones puntuales e intervalo de confianza al 95% de las elasticidades parciales de los distintos usos de suelo.

Fuente: elaboración propia.

Es de destacar que al considerar tanto de manera puntual como el límite superior del intervalo, la elasticidad parcial de la superficie agrícola (pea) es superior a la elasticidad parcial de la cubierta forestal (pef). Lo anterior indica que, manteniendo todo lo demás constante, se espera que incremento en el PIB sea mayor cuando se incrementa la superficie agrícola. Sin embargo, debe considerarse que el cambio en la superficie agrícola tiene un impacto indirecto, aunque presumiblemente más lento, en la escala de producción a través de la modificación en la fase de transición forestal. Debido a lo anterior, el impacto neto del incremento de la superficie agrícola resulta ambiguo sino se toman en cuenta dichas consideraciones de largo plazo.

Escala de producción

Los resultados obtenidos indican que la escala de producción (parámetro α_0 de la ecuación (1)) está determinada por los coeficientes β_0 , β_5 y β_6 . Lo cual indica que el parámetro mencionado varía de acuerdo a la etapa del proceso de transición forestal y la localización del país. En concreto se tiene que por cada combinación de etapa y continente (12 combinaciones en total) se definió un nivel distinto de escala de producción, siendo el nivel de referencia la etapa de pre transición y el continente África. Es decir que para los países con dichas características, la escala de producción está definida únicamente por β_0 . Es de

TRANSICIÓN DE ESTADÍSTICA A CIENCIA DE DATOS

destacar que, como lo indica el signo positivo de las variables fase y continente en todos sus niveles, el cambio en ellas tiende a aumentar la escala de producción. De manera concreta se observa que conforme se avanza en el proceso de transición forestal, la escala de producción aumenta. De manera similar, la escala de producción más baja corresponde al continente Africano, aumentando en América y Asia en el orden mencionado.

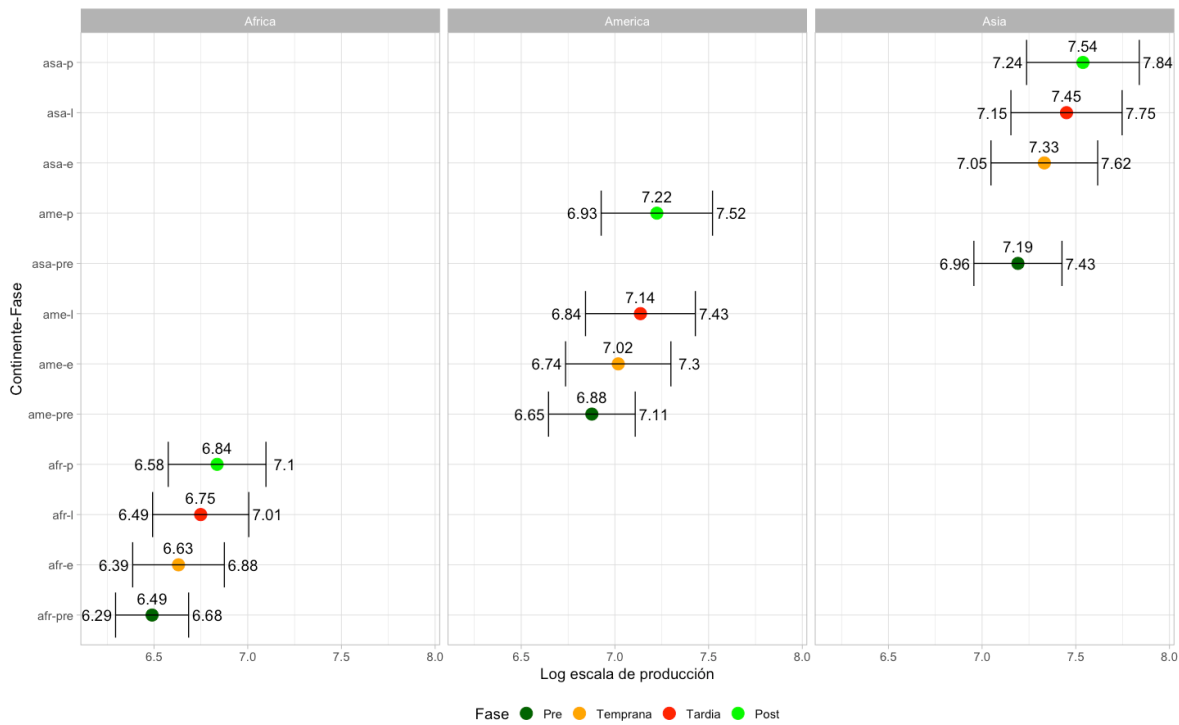


Figura 6. Estimación puntual e intervalo de confianza al 95% de la escala de producción por continente y fase.

Fuente: elaboración propia.

Los resultados numéricos de la estimación de escala de producción al combinar los factores fase y continente se muestran en la Figura 6. Como se puede anticipar, la escala de producción más baja corresponde al nivel de referencia (fase temprana en el continente africano), mientras que la más alta se observa en la etapa de postransición en el continente asiático. De manera similar se observa que, independientemente de la fase, la escala de producción más baja es en el continente africano, esta aumenta en América y alcanza los valores más altos en Asia. Otro patrón interesante a destacar es que la escala de producción sigue el mismo orden en función de la fase en todos los continentes. Es decir, es más baja

cuando el país se encuentra la fase de pre transición y aumenta conforme se avanza en el proceso de transición forestal, llegando a su punto más alto en la etapa de postransición. En dicha fase, la escala de producción es alrededor de 2.2 veces mayor con respecto a la pre transición en términos relativos en todos los continentes.

CONCLUSIONES

Este trabajo de investigación se planteo analizar, de manera empírica, la contribución del capital natural al bienestar en el sector de uso de suelo. Lo anterior se realizó a través de la detección de la existencia de servicios ecosistémicos mediante el uso de técnicas estadísticas. Dicho análisis se centró en los países de la región tropical, donde el balance de las necesidades de desarrollo económico y los objetivos de sostenibilidad resulta complejo.

El análisis reveló que los ecosistemas forestales juegan un papel fundamental en la creación de bienestar, ya que estos contribuyen a la creación de valor de manera directa (a través de la elasticidad parcial) e indirecta (a través de la escala de producción). Sin embargo, el análisis también reveló que existen incentivos privados para el cambio de uso de suelo o deforestación, debido a que la elasticidad parcial de la agricultura resulta superior a la del uso forestal.

Ante la contraposición de intereses colectivos e individuales, la regulación del sector en la región estudiada resulta imperativa. De igual modo se observa que una estrategia de conservación no se justifica mediante criterios de eficiencia económica. Por lo que futuras investigaciones deben enfocarse en evaluar como incide la inclusión de otros criterios en los resultados aquí obtenidos.

REFERENCIAS

- Carter, S., Herold, M., Rufino, M. C., Neumann, K., Kooistra, L., & Verchot, L. V. (2015). Mitigation of agriculture emissions in the tropics: comparing forest land-sparing options at the national level. *Biogeosciences*, 12. <https://doi.org/10.5194/bg-12-4809-2015>
- Costanza, R., de Groot, R., Braat, L., Kubiszewski, I., Fioramonti, L., Sutton, P., ... Grasso, M. (2017). Twenty years of ecosystem services: How far have we come and how far do we still need to go? *Ecosystem Services*, 28, 1–16. <https://doi.org/10.1016/j.ecoser.2017.09.008>

- Di Gregorio, M., Nurrochmat, D. R., Paavola, J., Sari, I. M., Fatorelli, L., Pramova, E., ... Kusumadewi, S. D. (2017). Climate policy integration in the land use sector: Mitigation, adaptation and sustainable development linkages. *Environmental Science & Policy*, 67, 35–43. <https://doi.org/10.1016/j.envsci.2016.11.004>
- FAO. (2014). *Building a Common Vision for Sustainable Food and Agriculture: Principles and Approaches*. Recuperado de <https://books.google.com.mx/books?id=6UTSrQEACAAJ>
- Godfray, H. C. J., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., ... Toulmin, C. (2010). Food Security: The Challenge of Feeding 9 Billion People. *Science*, 327(5967), 812. <https://doi.org/10.1126/science.1185383>
- Gujarati, D. N., Guerrero, D. G., & Medina, G. A. (2004). *Econometría*. McGraw-Hill.
- Hosonuma, N., Herold, M., de Sy, V., de Fries, R. S., Brockhaus, M., Verchot, L., ... Romijn, E. (2012). An assessment of deforestation and forest degradation drivers in developing countries. *Environmental Research Letters*, 7(4). <https://doi.org/10.1088/1748-9326/7/4/044009>
- IPCC. (2014). *IPCC, 2014: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 3–87). Geneva, Switzerland.
- Locatelli, B., Lavorel, S., Sloan, S., Tappeiner, U., & Geneletti, D. (2017). Characteristic trajectories of ecosystem services in mountains. *Frontiers in Ecology and the Environment*, 15(3). <https://doi.org/10.1002/fee.1470>
- MA. (2003). *Ecosystems and Human Well-being: A Framework for Assessment*. Recuperado de <https://books.google.com.mx/books?id=0JGLgfndz0YC>
- Raudsepp-Hearne, C., Peterson, G. D., Tengö, M., Bennett, E. M., Holland, T., Benessaiah, K., ... Pfeifer, L. (2010). Untangling the Environmentalist's Paradox: Why Is Human Well-being Increasing as Ecosystem Services Degrade? *BioScience*, 60(8), 576–589. <https://doi.org/10.1525/bio.2010.60.8.4>
- Reed, J., van Vianen, J., Foli, S., Clendenning, J., Yang, K., MacDonald, M., ... Sunderland, T. (2017). Trees for life: The ecosystem service contribution of trees to food production and livelihoods in the tropics. *Forest Policy and Economics*, 84, 62–71. <https://doi.org/10.1016/j.forpol.2017.01.012>

- Smith, P., Bustamante, M., Ahammad, H., Clark, H., Dong, H., Elsiddig, E. A., ... Tubiello, F. (2014). Agriculture, Forestry and Other Land Use (AFOLU). En *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Edenhofer, O., R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, A. Adler, I. Baum, S. Brunner, P. Eickemeier, B. Kriemann, J. Savolainen, S. Schlömer, C. von Stechow, T. Zwickel and J.C. Minx (eds.)]*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
- Varian, H. R. (2010). *Intermediate Microeconomics: A Modern Approach*. Recuperado de <https://books.google.com.mx/books?id=JMBfQgAACAAJ>
- World Bank. (2018a). Agriculture, forestry, and fishing, value added (current US\$) | Data. Recuperado el 14 de noviembre de 2018, de <https://data.worldbank.org/indicator/NV.AGR.TOTL.CD?view=chart>
- World Bank. (2018b). Forest area (% of land area) | Data. Recuperado el 24 de septiembre de 2018, de <https://data.worldbank.org/indicator/AG.LND.FRST.ZS?view=chart>
- World Bank. (2018c). Forest area (sq. km) | Data. Recuperado el 28 de septiembre de 2018, de <https://data.worldbank.org/indicator/AG.LND.FRST.K2?view=chart>
- World Bank. (2019). Agricultural land (sq. km) | Data. Recuperado el 3 de enero de 2019, de Agricultural land (sq. km) website: <https://data.worldbank.org/indicator/AG.LND.AGRI.K2?view=chart>

¿HIJOS POR ELECCIÓN? LA SITUACIÓN DE LA FECUNDIDAD EN MÉXICO, 2014.

Leonardo Ailines Genis

Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, UNAM

INTRODUCCIÓN

La fecundidad, junto con la mortalidad y la migración, forma parte de los tres componentes principales de la dinámica demográfica. La transición demográfica es un proceso de cambio demográfico en donde influyen aspectos sociales, económicos y culturales (Zabala de Cosío, 1992), y que ha explicado los cambios en la dinámica demográfica con el paso de los años, entre ellos los patrones de fecundidad.

Algunos estudios señalan que los niveles de fecundidad en América Latina han reducido notablemente en los últimos años a tal punto de presentar niveles cercanos a América del Norte y Oceanía; sin embargo, América Latina presenta importantes desigualdades en cada uno de sus países, registra grandes brechas en el estatus socioeconómico, el nivel escolar, el acceso a los servicios, entre otros (Rodríguez *et al.*, 2015). Otros autores mencionan que los cambios en los niveles de fecundidad se deben a los cambios sociales y económicos en torno a la mujer. Durante la mitad del siglo pasado Cuba reportó un promedio de 4.1 hijos por mujer (Chackiel, 2004), pero con la implementación de políticas públicas con relación al control de nacimientos y como uno de los países en América Latina que cuenta con una ley de despenalización del aborto (Alfonso, 2004), ha tenido como consecuencia una reducción notable en sus niveles de fecundidad, a tal punto de reportar 1.7 hijos por mujer en el año 2017 (Banco Mundial, 2019).

Mientras tanto, México comenzó a tener un avance importante en el proceso de transición demográfica durante el siglo XX. No obstante, no fue sino hasta la mitad de los años sesenta que comenzó a presentar un descenso importante en los niveles de fecundidad, esto debido a la difusión de las políticas públicas y programas de planificación familiar, así como algunos cambios sustanciales en la calidad de vida de los mexicanos (Morales *et al.*, 2016).

A pesar de ello, los cambios derivados de la transición demográfica son más significativos en ciertos grupos de la población y que, en algunas situaciones, están asociados a los niveles de vulnerabilidad, y en el caso de la fecundidad, a la exposición a un embarazo.

Es así como surge la idea central del presente trabajo de investigación: Identificar y analizar algunos factores sociodemográficos que pueden influir en el promedio de hijos nacidos vivos entre las mujeres mexicanas en edad fértil (de 15 a 49 años de edad).

La fecundidad y su relación con los factores sociodemográficos

Algunos investigadores indican que la edad es un factor importante que puede influir en el número de hijos que tiene una mujer. Welte (2012) menciona que el patrón de la fecundidad por edad sirve para derivar factores de ajuste aplicables a la fecundidad, que originalmente emanan solo del promedio de hijos nacidos vivos. En su trabajo muestra que dicho promedio aumenta conforme incrementa la edad de la madre, asimismo menciona que este promedio es más alto si la mujer fue madre durante la adolescencia.

Otro factor importante en el incremento de la fecundidad es la situación conyugal de las mujeres. Belmonte (2007) menciona en su trabajo que en el estado de Sonora las mujeres viudas, casadas y en unión libre son el grupo que promedian el mayor número de hijos, seguido de las mujeres divorciadas y finalmente las solteras. También, Curtis y Blanc (1997) han demostrado que las mujeres unidas tienen una mayor probabilidad de hacer un uso incorrecto de los métodos anticonceptivos en comparación con las solteras, ya que entre las mujeres solteras el uso efectivo de los métodos anticonceptivos está fuertemente relacionado con la prevención de un embarazo no planeado.

Asistir a una institución educativa es fundamental para disminuir la tasa de fecundidad, ya que en este lugar se brindan pláticas y talleres a los jóvenes sobre las distintas formas para prevenir un embarazo o una infección de transmisión sexual, a través del uso correcto de los métodos anticonceptivos (Stupp y Cáceres, 2001). Este último resultado indica que la escolaridad tiene un impacto significativo en el conocimiento y uso correcto de los anticonceptivos (y como consecuencia en la reducción en los niveles de fecundidad).

También, la relación entre el número de hijos nacidos vivos y el uso correcto de los métodos anticonceptivos tiene mayor fuerza cuando las mujeres tienen uno o dos hijos, ya que estas mujeres, junto con sus parejas, están más motivadas a evitar un embarazo no planeado a través del uso correcto de los anticonceptivos, en comparación con aquellas familias que aún no presentan descendencia o que alcanzaron una paridez de tres hijos o más (Curtis y Blanc, 1997).

La condición de actividad económica de las mujeres es considerada como otro factor que influye en la tasa de fecundidad. Algunos estudios indican que los cambios en las condiciones económicas de la población han generado transformaciones sociodemográficas de tal forma que dentro de los hogares es

frecuente encontrar a dos proveedores de ingresos (padre y madre) así como en aquellos donde la mujer es la única sostenedora de la familia (OIT, 2009).

Los niveles de fecundidad entre las mujeres indígenas son más elevados en comparación con el resto de la población, ya que forman parte de un grupo rezagado dentro de los programas sociales enfocados a la planificación familiar y acceso a los métodos anticonceptivos (Valdivieso, 2012). Szasz y Lerner (2010) mencionan en su trabajo que una de las causas por las que las mujeres indígenas no hayan escuchado hablar sobre los métodos anticonceptivos como medio de planificación familiar es debido al limitado acceso que tienen a estos, lo que permite pensar que dichas mujeres no tienen los elementos suficientes para regular su fecundidad ni de estar informadas sobre las características y consecuencias sobre el uso de métodos (Meneses *et al.*, 2017).

Otro factor que puede influir en la incidencia del embarazo es la localidad de residencia habitual de las mujeres. En las comunidades rurales se piensa que el embarazo está ligado a cuestiones culturales e ideológicas en las que se desenvuelven las mujeres, y que además se ha transmitido de generación en generación que su función es ser madres, por lo que el embarazo y el matrimonio forman parte de sus proyectos de vida (Tuñón y Eroza, 2001).

Finalmente se ha demostrado que los grupos socioeconómicamente más pobres no solo presentan la fecundidad más alta, sino que también son clasificados como un grupo vulnerable en materia de salud sexual y reproductiva (Di Cesare y Rodríguez, 2006), esto como consecuencia de un bajo acceso y disponibilidad de los métodos anticonceptivos, lo que pasa a ser un problema de necesidades insatisfechas (Schoumaker y Tabutin, 1999).

Por tal motivo, es necesario ubicar al embarazo como un proceso de cambio social y cultural que repercute en distintos grados de acuerdo al país y al contexto social en el que se ubica (Chalmers *et al.*, 2001).

Fuente de información y metodología

La fuente de información para esta investigación fue la Encuesta Nacional de la Dinámica Demográfica (ENADID) para su levantamiento del año 2014. Esta encuesta tiene como principal objetivo brindar información estadística relacionada con el nivel y el comportamiento de las componentes de la dinámica demográfica: fecundidad, mortalidad y migración, así como de brindar información con relación al tema de salud sexual y reproductiva. El tamaño de muestra fue de 101 mil 389 viviendas a nivel nacional, donde la unidad de observación fueron las mujeres de 15 a 54 años de edad residentes en las viviendas seleccionadas (INEGI, 2014).

Para determinar el número de hijos nacidos vivos (HNV) que han tenido las mujeres de 15 a 49 años de edad hasta el momento del levantamiento de la encuesta se consideró la siguiente pregunta dentro del cuestionario:

- En total, ¿cuántos hijos e hijas que nacieron vivos ha tenido usted?

Y las variables sociodemográficas seleccionadas para modelar el promedio de HNV fueron:

- Edad de la mujer al momento de la encuesta.
- Situación conyugal actual.
- Nivel de escolaridad.
- Conocimiento de los métodos anticonceptivos.
- Condición de actividad económica.
- Condición de habla de lengua indígena.
- Lugar de residencia habitual.²
- Estrato sociodemográfico.³

El modelo estadístico para explicar el promedio de HNV en las mujeres en edad fértil fue el Modelo de Regresión Poisson. Este modelo es considerado como referencia en estudios con variables de conteo (Cameron y Trivedi, 1998), el cual tiene como principal supuesto la equidispersión ($E[Y_i | x_1, \dots, x_n] = Var[Y_i | x_1, \dots, x_n]$), por lo que en algunas ocasiones los resultados que se obtienen suelen ser insatisfactorios si no se cumple este supuesto.

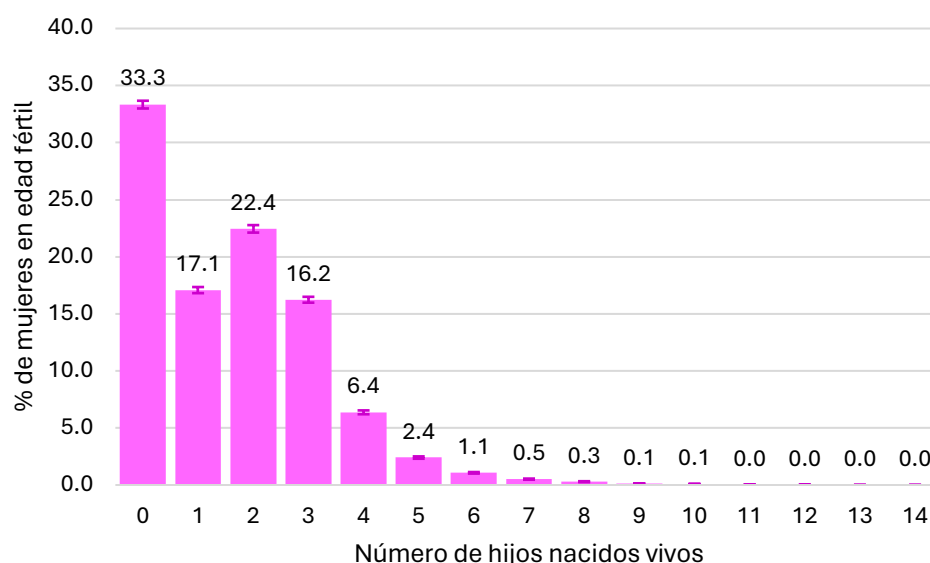
Análisis descriptivo de los factores que influyen en el número de HNV

De acuerdo con la información de la ENADID 2014, el total de mujeres en edad fértil encuestadas fue de 89,904 (se excluyeron para este análisis 12 mujeres de las cuales no se pudo obtener información puntual sobre su nivel de escolaridad). En la Gráfica 1 se muestra el porcentaje de mujeres en edad fértil según el número de HNV que han tenido hasta el momento del levantamiento de la encuesta.

² Urbano alto: 100,000 habitantes o más; Complemento urbano: entre 2,500 y 99,999 habitantes; Rural: menos de 2,500 habitantes.

³ Esta estratificación considera las características sociodemográficas de los habitantes en las viviendas seleccionadas, así como las características físicas y el equipamiento de estas, por medio de 34 indicadores clave.

Gráfica 1. Porcentaje de mujeres en edad fértil según número de hijos nacidos vivos, 2014.



Fuente: Estimaciones realizadas con base en el INEGI, Encuesta Nacional de la Dinámica Demográfica 2014.

En la gráfica anterior se muestra que el 33.3 por ciento de las mujeres en edad fértil no han tenido a su primer hijo nacido vivo hasta el momento del levantamiento de la encuesta; el 22.4 por ciento declaró haber tenido a dos hijos; el 17.1 por ciento mencionó que sólo han tenido a un hijo, y el 10.9 por ciento declaró tener a cuatro hijos o más. Estos resultados muestran que, en cuanto a preferencias reproductivas, las mujeres prefieren tener a dos hijos en vez de uno solo, sin embargo, aún prevalece un porcentaje importante de mujeres que no tienen hijos.

De acuerdo con la edad de las mujeres, el 87.3 por ciento de las mujeres de 15 a 19 años no han tenido a su primer HNV, mientras que sólo el 8.7 por ciento de las mujeres de 40 a 49 años se encuentran en esta situación. En el Cuadro 1 se puede observar que el número de HNV es mayor conforme incrementa la edad de la madre, de esta forma se puede suponer que la edad es un factor fundamental que explica los patrones de fecundidad.

Respecto a la situación conyugal, el 85.2 por ciento de las mujeres en edad fértil solteras no tienen hijos, mientras que entre las mujeres casadas o unidas este porcentaje disminuye a 8.5. En el Cuadro 1 se puede observar que las mujeres unidas y alguna vez unidas presentan un mayor número de hijos en comparación con las mujeres solteras, por lo que la situación conyugal puede ser un factor importante que puede determinar el número de HNV.

La ENADID 2014 determinó que el porcentaje más alto de mujeres sin hijos corresponde a aquellas que cuentan con un nivel de educación equivalente a la media superior o superior (46.3%), mientras que el porcentaje más bajo corresponde a las mujeres con sólo preescolar y primaria, seguidas de las mujeres sin instrucción alguna (13.1 y 14.6 por ciento, respectivamente). Un punto importante por analizar es el comportamiento del tamaño de la descendencia según la escolaridad de la madre ya que, como se muestra en el Cuadro 1, las mujeres con un mayor nivel de escolaridad prefieren familias pequeñas; mientras que, conforme reduce el nivel escolar, el tamaño de la descendencia tiende a aumentar.

Respecto al conocimiento sobre los métodos anticonceptivos, en el Cuadro 1 se puede observar que el 60.1 por ciento de las mujeres en edad fértil que no tienen hijos no tienen conocimiento sobre los distintos métodos anticonceptivos que existen; el 52.2 por ciento de ellas han escuchado hablar de ellos, sin embargo, no saben usarlos correctamente; finalmente solo el 31.4 por ciento conocen sobre los métodos anticonceptivos y su uso correcto. Es importante mencionar que, cuando se tiene uno o dos hijos, incrementa el conocimiento funcional de los métodos anticonceptivos, lo cual confirma que después del primer o el segundo hijo las parejas están más motivadas a prevenir un embarazo no planeado a través del uso correcto de los anticonceptivos.

De acuerdo con la condición de actividad, el porcentaje de mujeres en edad fértil económicamente activas sin hijos es de 32.2 por ciento, mientras que entre las no económicamente activas este porcentaje incrementa a 34.5. Es importante mencionar que el comportamiento es similar entre ambos grupos cuando se tiene una descendencia de uno a tres hijos; sin embargo, el porcentaje de mujeres con cuatro hijos o más es mayor entre las mujeres no económicamente activas (Véase Cuadro 1).

Según la condición de habla de lengua indígena de las mujeres, se muestra que existe una brecha significativa entre las mujeres sin hijos que hablan alguna lengua indígena y las que no hablan lengua indígena de aproximadamente ocho puntos porcentuales (Véase Cuadro 1). También se puede percibir que las mujeres que no hablan lengua indígena tienden a tener menos hijos en comparación con las mujeres que hablan algún dialecto, y cuya diferencia es más notable a partir del cuarto hijo.

Conforme al lugar de residencia habitual, el porcentaje de mujeres en edad fértil residentes en localidades urbanas sin hijos es de 35.8 por ciento, mientras que entre las mujeres residentes en las localidades rurales este porcentaje disminuye a 29.1 por ciento. De igual forma, se puede observar que el porcentaje de mujeres en edad fértil con cuatro hijos o más es más alto en las mujeres que viven en las localidades rurales (19.1%), en comparación con aquellas que viven en las zonas urbanas (7.4%) (Véase Cuadro 1).

TRANSICIÓN DE ESTADÍSTICA A CIENCIA DE DATOS

Finalmente, dentro de este análisis, se examinó el número de HNV de acuerdo al estrato sociodemográfico. El estrato sociodemográfico determina las características de los habitantes de las viviendas seleccionadas, así como de las características físicas y el equipamiento de las mismas, expresadas por medio de 34 indicadores clave (INEGI, 2014). Es así como el estrato sociodemográfico puede ser una variable que puede determinar algunas características de los niveles de pobreza dentro de una región.

En el Cuadro 1 se puede observar que la brecha que existe entre las mujeres sin hijos de acuerdo a su estrato sociodemográfico es de aproximadamente ocho puntos porcentuales. Si bien, la diferencia que existe entre ambos grupos en cuanto al número de hijos es mínima, se puede apreciar un aumento en el porcentaje de mujeres con cuatro o más hijos que pertenecen al estrato sociodemográfico bajo y medio bajo (13.3%) respecto a las que pertenecen a un nivel alto y medio alto (5.2%).

Cuadro 1. Porcentaje de mujeres en edad fértil según número de hijos nacidos vivos y características sociodemográficas de la madre, 2014.

Características sociodemográficas	Sin hijos		1 hijo		2 hijos		3 hijos		4 hijos o más		P-Value ^{4/}
	%	IC (90%)	%	IC (90%)	%	IC (90%)	%	IC (90%)	%	IC (90%)	
Edad											
De 15 a 19 años	87.3	(86.7, 87.9)	10.8	(10.3, 11.4)	1.7	(1.5, 2)	0.2	(0.1, 0.3)	-	-	0.0000
De 20 a 29 años	42.3	(41.6, 43)	26.8	(26.2, 27.4)	20.5	(20, 21.1)	7.8	(7.5, 8.1)	2.5	(2.3, 2.8)	
De 30 a 39 años	14.1	(13.6, 14.6)	16.3	(15.8, 16.8)	31.4	(30.8, 32)	24.4	(23.8, 25)	13.8	(13.3, 14.3)	
De 40 a 49 años	8.7	(8.2, 9.1)	10.6	(10.2, 11.1)	28.4	(27.7, 29.1)	27.6	(27, 28.3)	24.7	(24.1, 25.3)	
Situación conyugal											
Soltera	85.2	(84.7, 85.6)	9.5	(9.1, 9.9)	3.3	(3.1, 3.6)	1.4	(1.2, 1.5)	0.7	(0.6, 0.8)	0.0000
Viuda o separada	5.5	(5, 6)	25.5	(24.5, 26.6)	33.4	(32.3, 34.5)	21.0	(20.1, 22)	14.5	(13.8, 15.3)	
Casada o en unión libre	8.5	(8.3, 8.8)	20.0	(19.6, 20.4)	31.5	(31, 32)	23.9	(23.5, 24.3)	16.1	(15.7, 16.5)	
Nivel de escolaridad											
Media superior y superior	46.3	(45.8, 46.8)	19.8	(19.4, 20.2)	21.2	(20.7, 21.7)	10.2	(9.9, 10.5)	2.5	(2.4, 2.7)	0.0000
Secundaria	27.6	(27, 28.1)	17.0	(16.5, 17.4)	24.9	(24.3, 25.5)	19.8	(19.3, 20.2)	10.8	(10.4, 11.2)	
Preescolar o primaria	13.1	(12.5, 13.8)	11.5	(11, 12)	22.0	(21.3, 22.7)	25.0	(24.2, 25.7)	28.4	(27.6, 29.2)	
Sin instrucción	14.6	(12.7, 16.4)	9.0	(7.5, 10.5)	16.2	(14.3, 18.1)	18.1	(16.1, 20.1)	42.2	(39.8, 44.6)	
Conocimiento de los métodos anticonceptivos											
No conoce los métodos anticonceptivos	60.1	(57.3, 63)	10.0	(8.4, 11.6)	7.9	(6.5, 9.4)	5.2	(4.1, 6.3)	16.8	(14.8, 18.8)	0.0000
Sin conocimiento funcional ^{1/}	52.2	(50.9, 53.5)	12.5	(11.7, 13.3)	12.6	(11.8, 13.5)	10.4	(9.6, 11.1)	12.2	(11.4, 13.1)	

TRANSICIÓN DE ESTADÍSTICA A CIENCIA DE DATOS

Con conocimiento funcional ^{2/}	31. 4	(31, 31.7)	17. 6	(17.3, 17.8)	23. 5	(23.1, 23.8)	16. 9	(16.6, 17.2)	10. 7	(10.5, 10.9)	
Condición de actividad											
Económicamente activa	32. 2	(31.7, 32.7)	18. 7	(18.3, 19.1)	23. 8	(23.3, 24.3)	16. 0	(15.6, 16.3)	9.3	(9.1, 9.6)	0.0000
No económicamente activa	34. 5	(34, 34.9)	15. 5	(15.1, 15.8)	21. 1	(20.7, 21.5)	16. 5	(16.1, 16.9)	12. 5	(12.1, 12.8)	
Condición de habla de lengua indígena											
No habla lengua indígena	33. 8	(33.4, 34.1)	17. 3	(17, 17.5)	22. 7	(22.3, 23)	16. 2	(16, 16.5)	10. 0	(9.8, 10.2)	0.0000
Habla lengua indígena	25. 5	(24.2, 26.8)	13. 9	(13, 14.9)	18. 3	(17.2, 19.4)	16. 2	(15.1, 17.2)	26. 1	(24.8, 27.4)	
Lugar de residencia habitual ^{3/}											
Urbano alto	35. 8	(35.4, 36.3)	17. 8	(17.4, 18.1)	23. 6	(23.1, 24)	15. 4	(15.1, 15.8)	7.4	(7.2, 7.7)	0.0000
Complemento urbano	31. 3	(30.6, 31.9)	17. 1	(16.6, 17.7)	22. 2	(21.5, 22.8)	17. 7	(17.2, 18.2)	11. 7	(11.2, 12.1)	
Rural	29. 1	(28.4, 29.8)	15. 3	(14.7, 15.8)	19. 9	(19.2, 20.6)	16. 7	(16.1, 17.2)	19. 1	(18.5, 19.8)	
Estrato sociodemográfico											
Alto y medio alto	39. 1	(38.5, 39.7)	17. 7	(17.2, 18.2)	23. 6	(23.1, 24.2)	14. 3	(13.9, 14.8)	5.2	(4.9, 5.4)	0.0000
Bajo y medio bajo	30. 9	(30.5, 31.3)	16. 8	(16.5, 17.1)	21. 9	(21.5, 22.3)	17. 0	(16.7, 17.3)	13. 3	(13, 13.6)	

Nota: ^{1/} Conoce algún método anticonceptivo, pero no sabe usarlo correctamente.

^{2/} Conoce algún método anticonceptivo y sabe usarlo correctamente.

^{3/} Urbano alto: más de 100,000 habitantes; Complemento urbano: entre 2,500 y 99,999 habitantes; Rural: menos de 2,500 habitantes.

^{4/} Prueba Chi-Cuadrada de independencia.

Fuente: Estimaciones realizadas con base en el INEGI, Encuesta Nacional de la Dinámica Demográfica 2014.

Análisis multivariado de los factores que influyen en el promedio de HNV.

En el Cuadro 1 se puede observar que el número de HNV depende de cada uno de los factores sociodemográficos anteriormente analizados, ya que el p-value de la prueba Chi-Cuadrada de independencia es muy bajo (menor a 0.0000), por lo que se podrían considerar cada uno de estos en el modelo. Sin embargo, uno de los supuestos del modelo de regresión es la no correlación entre las variables independientes. La V de Cramer es un coeficiente basado en la prueba de la Chi-Cuadrada de independencia que mide la fuerza con la que están asociados dos campos categóricos. En el Cuadro 2 se muestra el valor de la V de Cramer para cada pareja de variables sociodemográficas.

Cuadro 2. Coeficiente V de Cramer de las variables sociodemográficas del Modelo de Regresión Poisson, 2014.

	Edad	Situación conyugal	Nivel de escolaridad	Conocimiento de los métodos anticonceptivos	Condición de actividad	Condición de habla de lengua indígena	Lugar de residencia habitual	Estrato sociodemográfico
Edad	1.0000							
Situación conyugal	0.4031	1.0000						
Nivel de escolaridad	0.1593	0.1785	1.0000					
Conocimiento de los métodos anticonceptivos	0.0744	0.0899	0.1868	1.0000				
Condición de actividad	0.2512	0.2037	0.1404	0.0510	1.0000			
Condición de habla de lengua indígena	0.0305	0.0562	0.2460	0.2179	0.0379	1.0000		
Lugar de residencia habitual	0.0292	0.0654	0.2141	0.1017	0.1728	0.2055	1.0000	
Estrato sociodemográfico	0.0552	0.0796	0.3115	0.0949	0.1125	0.1358	0.5063	1.0000

Fuente: Estimaciones realizadas con base en el INEGI, Encuesta Nacional de la Dinámica Demográfica 2014.

En el cuadro anterior se muestra que la asociación entre cada pareja de variables sociodemográficas es débil, a excepción de la edad y la situación conyugal, así como el lugar de residencia habitual y el estrato sociodemográfico, donde la asociación es moderada; sin embargo, en ninguno de los casos se muestra una asociación altamente significativa (mayor a 0.6), de manera que se pueden incorporar todas las variables analizadas al modelo. De esta manera, el Modelo de Regresión Poisson quedó definido como:

$$E[Y_i | x_{i,1}, x_{i,2}, \dots, x_{i,n}] = \mu_i = e^{\beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \dots + \beta_n x_{i,n}}$$

Donde:

Y_i representa el número de hijos nacidos vivos.

β_i son los coeficientes de regresión del modelo.

$x_{i,j}$ son las variables sociodemográficas anteriormente definidas que explican el promedio de hijos nacidos vivos (μ_i).

Como se mencionó anteriormente, el supuesto principal del Modelo de Regresión Poisson es la equidispersión en la variable de conteo. De acuerdo con la ENADID 2014, el promedio de HNV en las mujeres en edad fértil fue de 1.63 (IC(90%)=(1.62, 1.64)) con una varianza de 2.57 (IC(90%)=(2.53, 2.62)), por lo que, bajo una primera inspección, se puede suponer el supuesto de equidispersión; sin embargo, este resultado se debe validar con pruebas más robustas. Un modelo estadístico que permite evaluar si existe una sobredispersión en los datos es el Modelo de Regresión con distribución Binomial Negativa. En el Cuadro 3 se muestran los resultados obtenidos para ambos modelos de regresión (Poisson y Binomial Negativa).

Cuadro 3. Modelos de Regresión^{1/} Poisson y Binomial Negativa, 2014.

Variables	Distribución Poisson			Distribución Binomial Negativa		
	β	IC (90%)		β	IC (90%)	
Intercepto	-2.91	(-2.98, -2.83)	***	-2.91	(-2.98, -2.83)	***
Edad						
De 15 a 19 años (referencia)						
De 20 a 24 años	1.02	(0.97, 1.06)	***	1.02	(0.97, 1.06)	***
De 25 a 29 años	1.43	(1.39, 1.48)	***	1.43	(1.39, 1.48)	***
De 30 a 34 años	1.64	(1.6, 1.69)	***	1.64	(1.6, 1.69)	***
De 35 a 39 años	1.78	(1.73, 1.83)	***	1.78	(1.73, 1.83)	***
De 40 a 44 años	1.87	(1.83, 1.92)	***	1.87	(1.83, 1.92)	***
De 45 a 49 años	1.94	(1.89, 1.98)	***	1.94	(1.89, 1.98)	***
Situación conyugal						
Soltera (referencia)						
Viuda o separada	1.52	(1.48, 1.56)	***	1.52	(1.48, 1.56)	***
Casada o en unión libre	1.56	(1.52, 1.59)	***	1.56	(1.52, 1.59)	***
Nivel de escolaridad						
Media superior y superior (referencia)						
Secundaria	0.30	(0.29, 0.31)	***	0.30	(0.29, 0.31)	***
Preescolar o primaria	0.45	(0.43, 0.46)	***	0.45	(0.43, 0.46)	***
Sin instrucción	0.57	(0.54, 0.6)	***	0.57	(0.54, 0.6)	***
Conocimiento de los métodos anticonceptivos						
No conoce los métodos anticonceptivos (referencia)						
Sin conocimiento funcional ^{2/}	0.02	(-0.05, 0.08)		0.02	(-0.05, 0.08)	
Con conocimiento funcional ^{3/}	0.17	(0.11, 0.23)	***	0.17	(0.11, 0.23)	***
Condición de actividad						
Económicamente activa (referencia)						
No económicamente activa	0.05	(0.04, 0.06)	***	0.05	(0.04, 0.06)	***
Condición de habla de lengua indígena						
No habla lengua indígena (referencia)						
Habla lengua indígena	0.10	(0.08, 0.13)	***	0.10	(0.08, 0.13)	***
Lugar de residencia habitual						
Urbano alto (referencia)						
Complemento urbano	0.04	(0.03, 0.05)	***	0.04	(0.03, 0.05)	***
Rural	0.11	(0.1, 0.13)	***	0.11	(0.1, 0.13)	***
Estrato sociodemográfico						
Alto y medio alto (referencia)						
Bajo y medio bajo	0.07	(0.06, 0.09)	***	0.07	(0.06, 0.09)	***
Número de observaciones		89,904			89,904	

Número de parámetros	19	19
Devianza nula	159,629.9 con 89,903 grados de libertad	159,613.7 con 89,903 grados de libertad
Devianza residual	61,936.54 con 89,885 grados de libertad	61,929.52 con 89,885 grados de libertad
R ² Nagelkerke	0.6820	0.6387
Theta	-	14,287

Nota: ^{1/} Se ajustaron los coeficientes de regresión y sus varianzas debido al esquema de muestreo de la encuesta.

^{2/} Conoce algún método anticonceptivo, pero no sabe usarlo correctamente.

^{3/} Conoce algún método anticonceptivo y sabe usarlo correctamente.

*p<0.05; **p<0.01; ***p<0.001

Fuente: Estimaciones realizadas con base en el INEGI, Encuesta Nacional de la Dinámica Demográfica 2014.

Como se puede observar en el cuadro anterior, en ambos modelos se obtuvieron coeficientes de regresión similares (con un nivel de confianza del 90%), es decir, ambos modelos explican de manera similar el promedio de HNV. Para determinar la validez del supuesto de equidispersión se planteó la siguiente prueba de hipótesis:

$$H_0: \frac{1}{\theta} = 0 \quad vs \quad H_1: \frac{1}{\theta} > 0$$

Cameron y Trivedi (1998) proponen que el estadístico de prueba, basado en la log-verosimilitud de ambos modelos ($RV = -2[l_{poisson} - l_{BN}]$), se distribuye asintóticamente como una distribución Chi-Cuadrada con un grado de libertad, y se rechaza la hipótesis nula si el estadístico de prueba es mayor que $X^2_{1,1-\alpha}$. De esta forma, el estadístico de prueba arrojó un valor de $RV = -14.04$, que es menor a $X^2_{1,0.9} = 2.71$, por lo tanto, no se rechaza H_0 , es decir, los datos no presentaron problemas de sobredispersión, y en consecuencia se puede asumir que los datos cumplen el supuesto de equidispersión. En el Cuadro 4 se muestran los resultados obtenidos del Modelo de Regresión Poisson.

Cuadro 4. Modelo de Regresión Poisson^{1/} para explicar el promedio de hijos nacidos vivos en las mujeres en edad fértil, 2014.

Variables	Rate Ratio (RR)	IC (90%)	
Intercepto	0.05	(0.05, 0.06)	***
Edad			
De 15 a 19 años (referencia)			
De 20 a 24 años	2.76	(2.64, 2.89)	***
De 25 a 29 años	4.19	(4.01, 4.39)	***

TRANSICIÓN DE ESTADÍSTICA A CIENCIA DE DATOS

De 30 a 34 años	5.18	(4.95, 5.42)	***
De 35 a 39 años	5.93	(5.67, 6.21)	***
De 40 a 44 años	6.50	(6.21, 6.81)	***
De 45 a 49 años	6.95	(6.64, 7.28)	***
Situación conyugal			
Soltera (referencia)			
Viuda o separada	4.57	(4.39, 4.75)	***
Casada o en unión libre	4.74	(4.56, 4.92)	***
Nivel de escolaridad			
Media superior y superior (referencia)			
Secundaria	1.35	(1.33, 1.36)	***
Preescolar o primaria	1.56	(1.54, 1.58)	***
Sin instrucción	1.77	(1.71, 1.83)	***
Conocimiento de los métodos anticonceptivos			
No conoce los métodos anticonceptivos (referencia)			
Sin conocimiento funcional	1.02	(0.95, 1.09)	
Con conocimiento funcional	1.19	(1.11, 1.26)	***
Condición de actividad			
Económicamente activa (referencia)			
No económicamente activa	1.05	(1.04, 1.06)	***
Condición de habla de lengua indígena			
No habla lengua indígena (referencia)			
Habla lengua indígena	1.11	(1.09, 1.14)	***
Lugar de residencia habitual			
Urbano (referencia)			
Complemento urbano	1.04	(1.03, 1.06)	***
Rural	1.12	(1.1, 1.14)	***
Estrato sociodemográfico			
Alto y medio alto (referencia)			
Bajo y medio bajo	1.08	(1.06, 1.09)	***
Número de observaciones		89,904	
Número de parámetros		19	
Devianza nula	159,629.9 con 89,903 grados de libertad		
Devianza residual	61,936.54 con 89,885 grados de libertad		
R ² Nagelkerke		0.6820	
Prueba K-S sobre los residuos estandarizados		D = 0.1209, p-value < 2.2e-16	

Nota: ^{1/} Se ajustaron los coeficientes de regresión y sus varianzas debido al esquema de muestreo de la encuesta.

*p<0.05; **p<0.01; ***p<0.001

Fuente: Estimaciones realizadas con base en el INEGI, Encuesta Nacional de la Dinámica Demográfica 2014.

Es así que, en cuanto a la edad, el promedio de HNV en las mujeres de 20 a 24 años es 2.76 veces más alto que en las mujeres de 15 a 19 años; asimismo, el promedio de HNV entre las mujeres de 45 a 49 años es 6.95 veces más alto respecto al grupo de las adolescentes. Este resultado demuestra que a mayor edad de las mujeres mayor es el promedio de HNV, del mismo modo que lo menciona Welti (2012) en su trabajo.

De acuerdo a la situación conyugal, el promedio de HNV entre las mujeres alguna vez unidas es 4.57 veces más alto en comparación con las mujeres solteras. Este resultado es similar entre las mujeres unidas ($RR = 4.74$), por lo que se concluye que el promedio de HNV es más alto entre las mujeres unidas y alguna vez unidas, ya que estas tienen mayores probabilidades de concebir y formar una familia (Belmonte, 2007), asimismo las mujeres solteras buscan reducir las probabilidades de tener un embarazo no planeado a través del uso de métodos anticonceptivos (Curtis y Blanc, 1997).

También se puede observar que el promedio de HNV tiende a incrementar conforme reduce el nivel de escolaridad de las mujeres. De este modo, el promedio de HNV entre las mujeres que tienen un nivel de estudios equivalente a secundaria es 1.35 veces más alto respecto a aquellas que tienen algún grado en media superior o superior; mientras que entre las mujeres sin instrucción este promedio incrementa 1.77 veces. Es así como queda demostrado que a mayor nivel educativo en las mujeres es menor el promedio de HNV, por lo que asistir a una institución educativa es fundamental para reducir la tasa de fecundidad (Stupp y Cáceres, 2001).

El efecto del conocimiento de los métodos anticonceptivos sobre el promedio de HNV resultó ser no significativo entre las mujeres que no conocen los anticonceptivos y aquellas que no saben usarlos correctamente, es decir, el desconocimiento de los métodos anticonceptivos y la falla anticonceptiva derivada de su uso incorrecto tiene los mismos efectos en el tamaño de la descendencia. Sin embargo, el promedio de HNV entre las mujeres que tienen conocimiento funcional de los métodos anticonceptivos es 1.19 veces más alto en comparación con las mujeres que no han escuchado hablar de los métodos de planificación familiar, confirmando que las mujeres con uno o dos hijos están más motivadas a prevenir un embarazo no planeado a través del uso correcto de los métodos anticonceptivos (Curtis y Blanc, 1997). De acuerdo a la condición de actividad económica, el promedio de HNV entre las mujeres no económicamente activas es 1.05 veces más alto respecto a las mujeres económicamente activas. Los resultados muestran que las mujeres con la menor cantidad de hijos se dedican a alguna actividad económica, ya sea como apoyo de los ingresos en el hogar o en aquellos donde la mujer es la única sostenedora de la familia (OIT, 2009); mientras que, en las familias con una cantidad importante de hijos, la mujer se encarga de los cuidados del hogar, así como el cuidado de los hijos.

Según la condición de habla de lengua indígena de las mujeres en edad fértil, el promedio de HNV en las mujeres que hablan alguna lengua indígena o dialecto es 1.11 veces más alta en comparación con las mujeres que no hablan lengua indígena. Este resultado muestra la brecha que existe en la calidad y el acceso a los servicios de planificación familiar, ya que como mencionan algunos autores (Valdivieso, 2012;

Szasz y Lerner, 2010; Meneses *et al.*, 2017), este grupo se caracteriza por ser un sector de la población vulnerable en el tema de anticoncepción, debido a que dentro de estas comunidades el acceso a los anticonceptivos es limitado, y entre las mujeres que usan algún tipo de método la mayoría hacen uso de ellos por prescripción médica o preferencia de su pareja, lo que limita la oportunidad de una planificación familiar libre e informada (Szasz y Lerner, 2010).

Conforme al lugar de residencia habitual, el promedio de HNV en las mujeres que viven en localidades consideradas como complemento urbano es 1.04 veces más alta que el promedio de HNV en mujeres que viven en las zonas urbanas, mientras que dicho promedio en las localidades rurales es 1.12 veces más alta en comparación con el promedio en las localidades urbanas. Este resultado muestra que el promedio de HNV es más alto en las zonas rurales, posiblemente debido a que en dichas localidades se piensa que el embarazo está ligado a cuestiones culturales e ideológicas, donde el matrimonio y ser madres forman parte de los proyectos de vida de las mujeres (Tuñón y Eroza, 2001).

Conforme al estrato sociodemográfico, el promedio de HNV entre las mujeres dentro de los estratos bajos y medio bajos es 1.08 veces más alto en comparación con las mujeres en los estratos altos y medio altos. De esta manera se ha evidenciado la existencia de una relación directa entre la pobreza y la fecundidad en México: a mayor pobreza mayor fecundidad ya que, como lo mencionan Di Cesare y Rodríguez (2006), la población en los estratos más bajos es clasificada como un grupo vulnerable en materia de salud sexual y reproductiva a consecuencia de un bajo acceso y disponibilidad de los métodos anticonceptivos, lo que en la mayoría de las veces finaliza en un embarazo no planeado.

Finalmente, conforme al ajuste del modelo de regresión, la diferencia entre la devianza nula y la devianza residual es de 97,693.36 puntos, lo que sugiere que el ajuste es bueno. Asimismo, el porcentaje de la variación de los datos de la ENADID 2014 explicado por el Modelo de Regresión Poisson fue del 68.2%, de modo que, estadísticamente, es considerado un buen modelo. Sin embargo, el supuesto de normalidad sobre los residuos no se llegó a cumplir, ya que la prueba de Kolmogorov-Smirnov arrojó un estadístico de prueba de 0.1209 con un p-value muy bajo.

CONCLUSIONES

El Modelo de Regresión Poisson ha puesto en evidencia que la edad y la situación conyugal son las variables sociodemográficas más importantes para explicar el promedio de HNV en las mujeres en edad fértil en México. De esta forma, el promedio de HNV en una mujer casada o unida, de 30 años o más es mayor en comparación con aquellas mujeres solteras menores de 20 años, quienes buscan reducir las probabilidades de un embarazo no planeado a través del uso correcto de los métodos anticonceptivos.

En cuanto al nivel de escolaridad de las mujeres, aunque la universalidad de la educación a nivel secundaria sea un derecho en la actualidad, el análisis muestra que aún persiste un rezago educativo importante entre las mujeres, y que a su vez se ve reflejado en el número de HNV que han tenido. Por tal motivo es necesario seguir fomentando la educación a nivel media superior y superior, ya que permite a hombres y mujeres entender de mejor manera su cuerpo y su sexualidad, así como los riesgos que involucra tener relaciones sexuales sin protección alguna.

El uso correcto de métodos anticonceptivos es fundamental en la planificación familiar. Este análisis demostró que existe un mismo efecto sobre el tamaño de la descendencia en las mujeres que no conocen los métodos anticonceptivos, así como aquellas que no saben usarlos correctamente. También se ha puesto en evidencia que las mujeres tienen un mejor conocimiento funcional de los anticonceptivos sólo después del nacimiento del primer o del segundo hijo. El artículo 4 de la Constitución Mexicana establece que: “la planificación familiar es el derecho que toda persona tiene de decidir de manera libre, responsable e informada, el número y espaciamiento de sus hijos”, por tal motivo es necesario seguir fomentando esta garantía individual, así como disponer de la información y los medios necesarios para poder hacerlo.

Asimismo, los resultados del Modelo Poisson mostraron que el promedio de HNV es más alto entre las mujeres no económicamente activas. Cabe mencionar que, de acuerdo con las estimaciones del modelo, la condición de actividad económica es una variable altamente significativa, sin embargo, se podría plantear la hipótesis de que esta condición ya no será relevante para determinar el número de hijos en un futuro, debido a que han mejorado las condiciones económicas y sociales del país; asimismo, la participación de la mujer mexicana en el mercado laboral ha incrementado con el paso de los años.

Finalmente, este trabajo permitió demostrar que las mujeres que hablan alguna lengua indígena, viven en localidades rurales y cuyo estrato sociodemográfico es bajo, tienen una mayor propensión a un embarazo no planeado, misma que se refleja en la cantidad de hijos que tienen, esto como resultado de una deficiente difusión de los métodos anticonceptivos, el bajo acceso a los servicios de salud en las localidades rurales, la falta de personal capacitado para transmitir en lengua indígena el uso correcto de los anticonceptivos

como medio de planificación familiar, así como medio para prevenir una infección de transmisión sexual, entre otros.

Es importante mencionar que en este trabajo se analizaron algunos factores que pueden influir en el promedio de HNV en el país, no obstante, la fecundidad como parte de la dinámica demográfica es considerada como un fenómeno multicausal, en donde pueden intervenir factores que en este trabajo no se hayan mencionado, por lo que se sugiere analizar otras fuentes de información, así como otras metodologías que puedan explicar de manera más certera dicho fenómeno.

REFERENCIAS

- Alfonso, Juan Carlos (2004). “*El descenso de la fecundidad en Cuba: de la primera a la segunda transición demográfica*”. En: *Revista Cubana de Salud Pública*, vol. 32, núm. 1, Centros de Estudios de Población y Desarrollo. Pág. 1-19.
- Banco Mundial (2019). Tasa de fertilidad, total (nacimientos por cada mujer). Fecha de consulta: 27/04/2019. Disponible en: <https://datos.bancomundial.org/indicador/SP.DYN.TFRT.IN>
- Belmonte, Gregorio (2007). *Construcción de masculinidades y su incidencia en el embarazo adolescente del Valle del Yaqui 2000-2004*. Tesis de Doctorado. Instituto de Investigaciones Histórico Sociales. Universidad Veracruzana. Fecha de consulta: 11/04/2019. Disponible en: <https://cdigital.uv.mx/handle/123456789/41278>
- Cameron, A.C. y Trivedi, P.K. (1998). *Regression Analysis of Count Data*. Econometric Society Monographs. Cambridge University Press.
- Chackiel, Juan (2004). “*La transición de la fecundidad en América Latina*”. En: *Notas de Población*, vol. 10, núm. 41, Colegio de México. Pág. 9-58.
- Chalmers, H.; Ingham, R.; Cleland, J.; Castle, S.; Stone, N.; Aggleton, P.; Partridge, R. (2001). *Dynamic contextual analysis of young peoples sexual health*. Faculty of Social Sciences. University of Southampton. Fecha de consulta: 11/04/2019. Disponible en: <https://www.gov.uk/dfid-research-outputs/dynamic-contextual-analysis-of-young-people-s-sexual-health>
- Curtis, Sian L. y Blanc, Ann K. (1997). *Determinants of contraceptive failure, switching and discontinuation: An analysis of dhs contraceptive histories*. En dhs Analytical Reports, num. 6, Calverton, Maryland. Fecha de consulta: 11/04/2019. Disponible en: <https://www.dhsprogram.com/pubs/pdf/AR6/AR6.pdf>

- Di Cesare, Mariachiara; Rodríguez, Jorge (2006). “Análisis micro de los determinantes de la fecundidad adolescente: en el caso de Brasil y Colombia”. En: *Papeles de Población*, Año 12, No. 18, abril-junio. Pág. 107-140.
- INEGI [Instituto Nacional de Estadística y Geografía] (2014). *Encuesta Nacional de la Dinámica Demográfica, 2014*. Fecha de consulta: 01/04/2019. Disponible en: <https://www.inegi.org.mx/programas/enadid/2014/>
- Meneses, Eloina; Ailines, Leonardo; Hernández, María Felipa (2017). “Factores que influyen en el conocimiento funcional de métodos anticonceptivos”. En: *La Situación demográfica de México 2017*, CONAPO, México. Pág. 65-92.
- Morales, Ángel Jair; Silva, Eliud; Ordorica-Mellado, Manuel (2016). “Pronósticos de la fecundidad en México: una aplicación de los modelos multivariados de series de tiempo”. En: *Papeles de Población*, Año 22, No. 89, Universidad Autónoma del Estado de México. Pág. 99-131.
- OIT [Organización Internacional del Trabajo] (2009). *Trabajo y Familia: hacia nuevas formas de conciliación con corresponsabilidad social*. Programa de las Naciones Unidas para el Desarrollo (PNUD). Fecha de consulta: 11/04/2019. Disponible en: http://www.ilo.org/wcmsp5/groups/public/---dgreports/---gender/documents/publication/wcms_111376.pdf
- Rodríguez, Grisell; Molina, Matilde; y Quintana, Livia (2015). “Fecundidad. Estudio y comportamiento. América Latina y Cuba”. En: *Novedades en población*, vol.11, núm. 21, Centro de Estudios Demográficos. Pág. 65-77.
- Schoumaker, Bruno y Tabutin, Dominique (1999). *Relations entre pauvreté et fécondité dans les pays du Sud. Connaissances, méthodologie et illustrations*. Document de Travail, No. 2, Département des sciences de la population et du développement, Louvain-la-Neuve, Université Catholique de Louvain. Fecha de consulta: 11/04/2019. Disponible en: <https://cdn.uclouvain.be/public/Exports%20reddot/sped/documents/WP2.pdf>
- Stupp, Paul; Cáceres, José Mario (2001). *The relationship between age at completion of schooling and age at first birth in El Salvador*. Ponencia presentada en la XXIV Conferencia, General de Población de la IUSSP, del 17 al 23 de agosto, Salvador, Bahía.
- Szasz, Ivonne y Lerner, Susana (2010). “Salud Reproductiva y Desigualdades en la Población”. En: *Los grandes problemas de México, Tomo I. Población*, Colegio de México. Pág. 214-251.

- Tuñon, Esperanza y Eroza Enrique (2001). “*Género y sexualidad adolescente. La búsqueda de un conocimiento huido*”. En: *Estudios sociológicos*, enero-abril, vol. XIX, núm 001. Colegio de México. Pág. 209-226.
- Valdiviezo, Noé (2012). *Factores asociados al embarazo en adolescentes indígenas de la región de los valles centrales de Oaxaca, 2014*. Tesis de Maestría. El Colegio de la Frontera Norte. Tijuana, B.C., México. Fecha de consulta: 11/04/2019. Disponible en: <https://www.colef.mx/posgrado/wp-content/uploads/2014/03/TESIS-Valdiviezo-Villanueva-No%C3%A9.pdf>
- Walti, Carlos (2012). “*Análisis de la fecundidad en México con los datos del Censo de Población y Vivienda 2010*”. En: *Papeles de Población*, vol. 18, núm. 73, Universidad Autónoma del Estado de México. Pág. 45-76.
- Zabala de Cosío, María Eugenia (1992). “*La transición demográfica en América Latina y en Europa*”. En: *Notas de Población*, vol. 20, núm. 56, Colegio de México. Pág. 11-32.

TIPOS DE VIOLENCIA QUE EXPERIMENTAN LOS ESTUDIANTES UNIVERSITARIOS. UN ANÁLISIS ESTADÍSTICO CON (SPSS)

Jeysira Jacqueline Dorantes Carrión

Universidad Veracruzana

jedorantes@uv.mx

RESUMEN

En el presente artículo se exponen datos sobre ocho tipos de violencia que se ejercen en el interior de la Universidad Veracruzana (UV): física, económica, verbal, docente, de género, sexual, psicológica o emocional, y el *cyberbullying*. El objetivo es describir las representaciones sociales que los estudiantes han construido acerca de estos tipos de violencia, a partir de su experiencia. La investigación siguió una metodología cuantitativa, con la encuesta como técnica de recolección de datos; el instrumento de medición fueron ocho preguntas abiertas y la muestra estuvo constituida por 1931 estudiantes. Los resultados arrojaron un conjunto de palabras, de las cuales, el núcleo central de las representaciones sociales de los tipos de violencia son las palabras golpes; falta de dinero; insultos, ofender y groserías; daño psicológico y emocional; abuso de autoridad; discriminación y humillación; forzar a tener relaciones sexuales; agresión en Internet; mientras que los elementos periféricos en torno al núcleo son maltrato a tu persona; explotación laboral; gritos, insultos, ofensas y groserías; maestro violento; machismo; acoso sexual; acoso virtual, entre otros. En conclusión, se asume que existe un conjunto de prácticas violentas que ejercen y padecen, de manera simultánea, los jóvenes universitarios, y que es necesario atenderlas de manera focalizada.

Palabras clave: entorno universitario, estudiantes, manifestaciones violentas, representaciones sociales, tipos de violencia.

INTRODUCCIÓN

Este documento se desprende de la investigación denominada “Representaciones sociales de género entre estudiantes de la Universidad Veracruzana”⁴, donde se examinó la violencia que viven los estudiantes en

⁴ La investigación responde al proyecto Conacyt núm. 197716.

el entorno universitario. Se llevó a cabo con un grupo de alumnos de las cinco regiones que conforman la Universidad Veracruzana (UV).

Para efectos de este apartado, se aborda como objeto de estudio a la violencia y su tipología (Arendt, 2005; Bringiotti, Krynveniuk y Lassi, 2005; Buquet, 2013), dado que forma parte de un fenómeno social que afecta a los estudiantes de los diferentes niveles educativos, incluyendo a la universidad.

De acuerdo con la Organización Mundial de la Salud (OMS, 2020b) la violencia es:

El uso intencional de la fuerza física, amenazas contra uno mismo, otra persona, un grupo o una comunidad que tiene como consecuencia o es muy probable que tenga como consecuencia un traumatismo, daños psicológicos, problemas de desarrollo o la muerte (párr. 1).

Sus efectos son experimentados y sufridos en espacios como las aulas, cubículos, laboratorios, pasillos, bibliotecas, áreas verdes y explanadas, así como fuera de los espacios escolares. La violencia es un problema mayor, considerando que se manifiesta de manera: a) interna, permea los espacios educativos en donde conviven profesores, autoridades, estudiantes y otros actores educativos; y b) externa, en el ámbito familiar, la comunidad, la política, el gobierno y en la sociedad. Toda la estructura de nuestra personalidad está vertebrada por ésta, y experimentamos cierto reparo o repugnancia, o cuando menos aversión, ante el uso de violencia, y este proceso se ha ido desarrollando a lo largo de los años (Kaplan, 2016, p. 215).

Lo anterior trae consigo efectos negativos para las personas en general, así como para los estudiantes, dado que “afecta de manera determinante a los alumnos, en sus estudios y en su aprovechamiento escolar” (Furlán y Spitzer, 2013). Además, la OMS (2020b) declaró que “la violencia juvenil tiene repercusiones graves, que a menudo perduran toda la vida, en el funcionamiento físico, psicológico y social de una persona” (párr. 5). Lo grave del tema es que muchos jóvenes no hablan de la violencia que experimentan, lo cual los posiciona en una situación de vulnerabilidad; por lo general se callan, ocultan lo que les está sucediendo, muchas veces por miedo a ser reprimidos por quien los agrede, o bien, porque desconocen a dónde buscar ayuda o con quién solicitar orientación.

Lamentablemente, en la UV aún no se cuenta con un diagnóstico amplio que muestre la incidencia de los tipos de violencia detectados como más frecuentes en el espacio universitario. Casillas, Dorantes y Ortiz (2017), nos explican “aparentemente las universidades son inocuas a los problemas sociales; se representan como templos de saber, espacios privilegiado donde se produce y difunde el conocimiento. Sin embargo, en sus aulas se expresan los principales problemas sociales contemporáneos y se ponen en evidencia las profundas desigualdades que caracterizan la experiencia escolar de los estudiantes” (p. 7).

Por su parte los autores Casillas, Dorantes, Guadarrama, Castro y Morales (2017), precisan que “si una tercera parte de los estudiantes reporta haber sido víctima de violencia física con padres y hermanos, estaríamos hablando de unos 20 mil estudiantes que tiene inscrita en su historia este tipo de violencia. 9 o 10 mil han sufrido violencia económica en sus familias. Si dos terceras partes de la población ha sufrido violencia verbal en sus hogares, en las escuelas previas e incluso en la universidad, estamos hablando de un universo cercano a los 40 mil estudiantes. Así podríamos seguir enumerando y considerando el volumen y el alcance de los actos violentos en la vida de los estudiantes de la UV” (p. 76).

En ese sentido, lo que se buscó en este estudio es conocer la realidad sobre la tipología predominante de la violencia en el interior de la universidad, así como describir las representaciones sociales construidas por los estudiantes en torno a esta.

Para ello, se retomó como referente teórico la Teoría de las Representaciones Sociales (TRS) de Moscovici (1961), cuyo origen es la psicología social y la sociología del conocimiento (Banchs, 2000). Según Moscovici (1961), la representación social es “una organización de imágenes y de lenguaje porque recorta y simboliza actos y situaciones que son o se convierten en comunes, (...) [por lo que] una representación siempre es de alguien, así como es una representación de una cosa” (pp. 16-17). Es decir, se trata de un conocimiento que se desarrolla a partir del sentido común de las personas, donde interviene lo psicológico y social, y se produce mediante sus interacciones en la vida cotidiana.

Por su parte, para Jodelet y Guerrero (2000) las representaciones sociales son:

Conciernen en el pensamiento de sentido común, que se pone a disposición en la experiencia cotidiana; son programas de percepción, construcciones con estatus de teoría ingenua, que sirven de guía para la acción e instrumento de lectura de la realidad; sistemas de significaciones que permiten interpretar el curso de acontecimientos y las relaciones sociales; que expresan la relación que los individuos y los grupos mantienen con el mundo y los otros; que son forjadas en la interacción y el contacto con los discursos que circulan en el espacio público; que están inscritas en el lenguaje y las prácticas; y que funcionan como un lenguaje en razón de su función simbólica y de los marcos que proporcionan para codificar y categorizar lo que compone el universo de la vida (p.10).

Como puede apreciarse, una representación social permite captar la realidad a partir de la mirada de los individuos y los grupos, considerando que posee características esenciales para ello. Como lo menciona Álvarez (2004):

Las representaciones sociales están impregnadas de los fenómenos culturales como son los políticos, religiosos, etc. Fenómenos a los que se supone que la psicología social debe aportar conocimientos para explicarlos...las representaciones sociales son elaboradas por un comité de decisión colectiva, donde los miembros del mismo elaboran sus votos y pueden expresar un rango amplio de opiniones, donde cada uno de los miembros sabe que los otros pueden votar y que pueden cambiar sus pensamientos y combinar sus opiniones, y que el resultado final vendría a ser el mutuo esfuerzo de todos los participantes expresado en un sentido de compromiso (p. 37).

Por su parte Como lo mencionó Jodelet (1986) las representaciones sociales:

- Invariablemente representa un objeto.
- Posee carácter de imagen y la propiedad de intercambiar percepción, pensamiento y concepto.
- Tiene una naturaleza simbólica y significativa.
- Posee propiedades constructivas.
- Está dotada de un carácter autónomo e innovador (como se citó en Villarroel, 2007, p. 441).

En este caso, se puede lograr explicar que las propuestas y opiniones individuales se encuentran vinculadas con las opiniones del grupo, y que es aceptable y comprendida por todos los que integran a dicho grupo; sin embargo, se trata de identificar dos componentes fundamentales de la representación social: el núcleo central y los elementos periféricos, es decir, la palabra que representa al objeto en estudio (tipos de violencia), y otras palabras que le otorgan un sentido, significación, relación y razón de ser. De acuerdo con Abric (1994), “toda representación social está organizada alrededor de un núcleo central” (p. 20) y de elementos periféricos que se encuentran jerarquizados, los cuales “desempeñan un papel importante en la concreción del significado de la representación, más distantes del núcleo, ilustran, aclaran y justifican esta significación” (p. 30).

Al respecto, Jodelet y Guerrero (2000) señalaron que dichos elementos –el núcleo central y los elementos periféricos– constituyen lo esencial del contenido de la representación, presentan su lado más accesible, pero también lo más vivo y concreto de esta. Asimismo, para Flament (1994), los elementos periféricos “son un conjunto organizado de cogniciones relativas a un objeto, compartidas por los miembros de una población homogénea respecto a ese objeto” (p. 37).

En otras palabras, el individuo instauro, organiza y jerarquiza el contenido de la representación social, le asigna un valor y le atribuye un significado; la representación abarca informaciones retenidas, seleccionadas e interpretadas; juicios formulados, estereotipos o creencias al respecto del objeto y su entorno. En ese sentido, el propósito es describir las representaciones sociales que los estudiantes

universitarios han construido sobre los tipos de violencia, a partir del análisis del núcleo central y de los elementos periféricos que se organizan a su alrededor, y que están en relación directa con ellos.

En esa dirección, se muestran los tipos de violencia analizados en este artículo, y que el portal Significados.com (2019) definió de la siguiente manera:

1. Violencia psicológica o emocional: Es aquella que produce daños afectivos, morales y psicológicos sobre la persona sin que medie la fuerza física, lo que hace mermar su autoestima (párr. 4).
2. Violencia física: Se usa la fuerza física, causa dolor, daño, heridas y hasta la privación de la vida (párr. 5).
3. Violencia sexual: Es aquella que denigra la integridad sexual de una persona, sea que haya contacto genital o no. Se entiende que la integridad ha sido vulnerada toda vez que la víctima no da su consentimiento expreso (párr. 6).
4. Violencia económica: Se ejerce contra una persona, vulnerando sus derechos económicos por medio de la perturbación de la posesión o propiedad de sus bienes, así como por medio de la sustracción, destrucción deliberada y retención de bienes, herramientas de trabajo, documentos personales, bienes y valores (párr. 7).
5. Violencia verbal: Agresión por medio de insultos y descalificaciones (párr. 4).
6. Violencia de género: Ejercida sobre la persona a causa de su género. La mayor parte de las víctimas de este tipo de violencia son las mujeres, cuyo maltrato está absolutamente naturalizado en la cultura. Entre las víctimas cuentan también los miembros de la comunidad LGBT. La violencia de género puede ir desde la descalificación hasta el feminicidio, en cuyo caso califica como crimen de odio (párr. 13).
7. *Cyberbullying*: Acoso por medio de las redes sociales, en el que se alcanzan altos niveles de violencia psicológica, potenciado por el anonimato del Internet (párr. 15).
8. Violencia docente: Aquella que ejercen los profesores hacia los alumnos, ya sea por cualquier condición física, intelectual, emocional, social, etc. Forma de mantener el poder y el orden dentro de un grupo determinado (Conceptodefinición.de, 2016, párrs. 1-3).

Particularmente Dorantes (2017), manifiesta que “en la UV no deberían de existir ninguno de los tipos de violencia, su presencia es inaceptable ya que la violencia refiere a actos con mala intención o ejecutados con saña”, ... Sin embargo, afirma que muchos de los estudiantes universitarios experimentan de manera

simultánea, estos tipos de violencia enunciados, que viven o sufren, particularmente en todas las regiones y áreas de conocimiento de la universidad” (p. 107).

El estudio permitió responder a la pregunta ¿cuáles son las representaciones sociales que los estudiantes universitarios han construido sobre los tipos de violencia (violencia, física, económica, verbal, psicológica o emocional, docente, de género, ciberbullying y sexual)? Al mismo tiempo, nos invita a reflexionar sobre el problema de la violencia, con el fin de contribuir a la construcción de una mejor sociedad. El proceso civilizatorio requiere de un alto grado de autodisciplina por parte de los miembros de la sociedad y una significativa pacificación interna de la sociedad moderna (Kaplan, 2016, p. 216).

METODOLOGÍA

Como abordaje metodológico se empleó el enfoque cuantitativo (Combessie, 2003; Taylor y Bogdan, 2006; Marradi, Archiento y Piovani, 2007). Para determinar la población de estudio se contempló el diseño muestral estratificado. Se realizó una selección aleatoria de las facultades, Universidad Veracruzana Intercultural (UVI) y posgrados a encuestar dentro de las cinco regiones que conforman la UV: Xalapa, Veracruz, Poza Rica, Orizaba y Coatzacoalcos; en seis áreas de conocimiento: Artes, Biológicas-Agropecuarias, Económico-Administrativa, Humanidades, Ciencias de la Salud y Técnica; donde cada elemento de la población objetivo tuvo la misma probabilidad de ser seleccionado. La fórmula de la muestra que se empleó es la siguiente:

$$n = \frac{\sum_{i=1}^L N_i^2 p_i q_i / w_i}{N^2 D + \sum_{i=1}^L N_i p_i q_i}$$

$$D = \frac{B^2}{Z_{\alpha/2}^2} \quad \text{y} \quad w_i = \frac{n_i}{n}$$

El tamaño de la muestra es de 1931 estudiantes y su representatividad es por región. Véase la siguiente tabla.

Tabla 1. Población y muestra

Región	Técnica	Humanidades	Económico-Administrativo	Ciencias de la salud	Biológico-Agropecuaria	Artes	SEA*	UVI	Posgrado	Total
Xalapa	156	185	183	120	36	39	78	0	43	840
Veracruz-Boca del Río	76	37	82	116	21	0	38	0	17	387
Poza Rica-Tuxpan	79	28	16	48	26	0	24	14	7	242
Córdoba-Orizaba	79	0	48	51	18	0	48	12	6	262
Coatzacoalcos-Minatitlán	55	11	40	38	7	0	29	12	8	200
Total	445	261	369	373	108	39	217	38	81	1931

*Sistema de enseñanza abierta.

Fuente: Elaboración propia.

Instrumento de medición

En el estudio “Representaciones sociales de género entre estudiantes de la Universidad Veracruzana”, proyecto Conacyt núm. 197716, se empleó el cuestionario como instrumento de medición, constituido por 60 preguntas. “El cuestionario, es el instrumento de recolección de datos propiamente dicho, utilizado en la técnica de encuestas y en la investigación (Marradi, Archiento y Piovani, 2007, p. 211); y “permitió identificar las respuestas y poner de manifiesto los factores explicativos o discriminantes de una población, identificar y situar posiciones, respecto a sus ejes explicativos” (Abrie, 1994, p. 56). El modo de aplicación del cuestionario fue cara a cara, lo cual permitió una mejor interacción con los informantes.

Una vez que se efectuó el trabajo de campo (aplicación del cuestionario), se capturó la información en Excel para formar una base de datos; posteriormente, se empleó el *software* Statistical Package for the Social Sciences (SPSS), que permitió organizar los datos para emprender el proceso de análisis de la información. Dicho paquete estadístico denominado permitió emprender el proceso de análisis procedente de la encuesta aplicada.

Sin embargo, para efectos de este escrito, solamente se tomaron en consideración ocho preguntas abiertas de dicho cuestionario, referentes a cada uno de los tipos de violencia. Se le pidió a cada estudiante que enunciara cinco palabras que asociara con los tipos de violencia anteriormente descritos: física, económica, verbal, psicológica o emocional, docente, de género, sexual y el *cyberbullying*; por ejemplo, ¿cuáles son las cinco palabras que asocias cuando hablamos de violencia sexual? Con base en las repuestas, se

agruparon los términos, permitiendo identificar el núcleo central de la representación social de la tipología de la violencia y sus elementos periféricos.

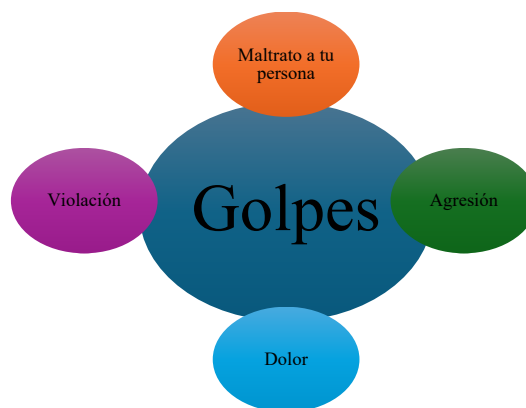
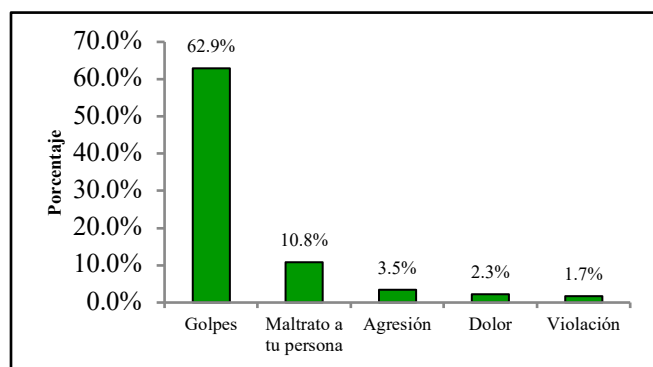
RESULTADOS

A continuación, se exponen los ocho tipos de violencia predominantes en el entorno universitario, donde se describen las respuestas obtenidas de los 1931 estudiantes de la UV. Es necesario destacar que los valores no suman 100% debido a que solamente se presenta una parte del concentrado general.

Violencia física

Al analizar la pregunta ¿Cuáles son las cinco palabras que asocias cuando hablamos de violencia física?, se encontró que el 62.9% de los participantes mencionó la palabra golpes; el 10.8% (179 estudiantes) la asocia con maltrato a tu persona; la tercera asociación es agresión, con el 3.5%; seguida de dolor, con un 2.3%; la última asociación fue violación, con el 1.7%, término que tuvo menor impacto en la representación, pero no por ello menos significativo.

Gráfica 1. Principales palabras que asocian cuando se habla de violencia física

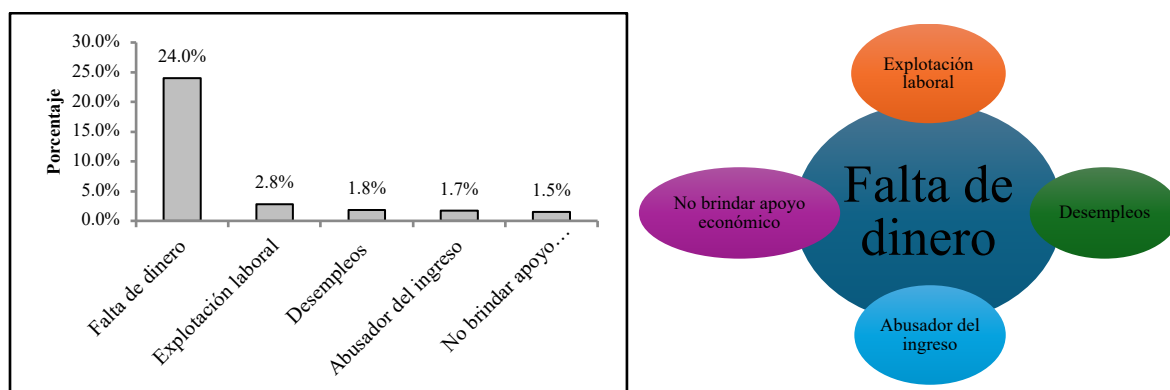


Se puede visualizar que la representación social de la violencia física se compone de la manifestación “golpes” como núcleo central y se articula con otros actos violentos, denominados elementos periféricos: “maltrato a tu persona, agresión, dolor y violación”. Estas expresiones son producto de un proceso de jerarquización y ordenamiento de la representación social.

Violencia económica

El resultado respecto a este tipo de violencia arroja como núcleo central de la representación social a la “falta de dinero”, con el 24.0%, es decir, la cuarta parte de los encuestados así lo percibe; los elementos periféricos son “explotación laboral” (2.8%), “desempleos” (1.8%), “abusador del ingreso” (1.7%) y “no brindar apoyo económico” (1.5%).

Gráfica 2. Principales palabras que asocian cuando se habla de violencia económica



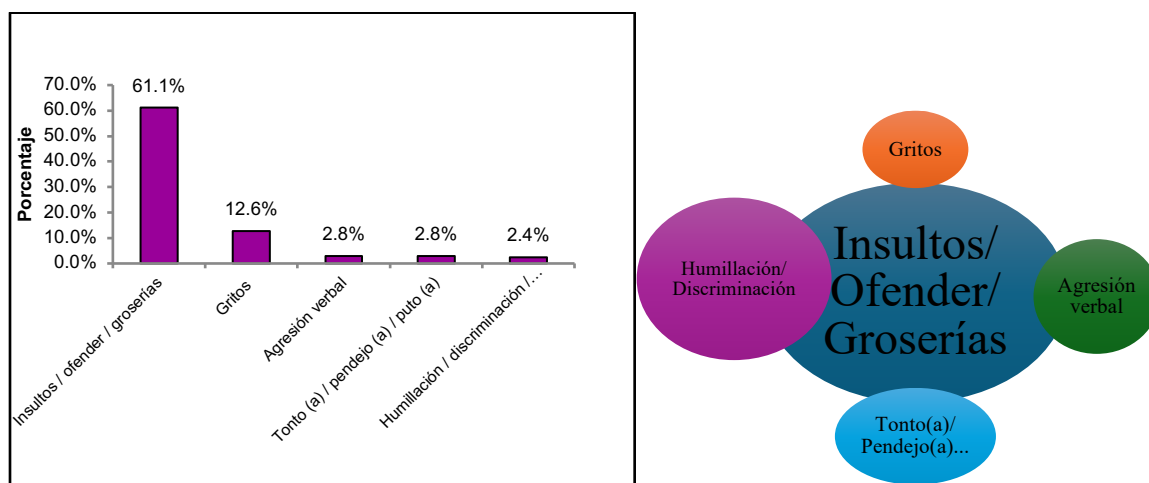
Fuente: Elaboración propia.

Al observar el esquema anterior, damos cuenta de que el dato duro es la “falta de dinero”, lo que habla de que un estudiante universitario se siente vulnerable ante esta violencia cuando no tiene dinero o es escaso.

Violencia verbal

Ante este tipo de violencia, los encuestados señalaron que se vincula con “insultos, ofender y groserías” (61.1%); “gritos” (12.6%); “agresión verbal” (2.8%); las expresiones “tonto(a), pendejo(a), puto(a)” (2.8%); y “humillación, discriminación y denigración” (2.4%).

Gráfica 3. Principales palabras que asocian cuando se habla de violencia verbal



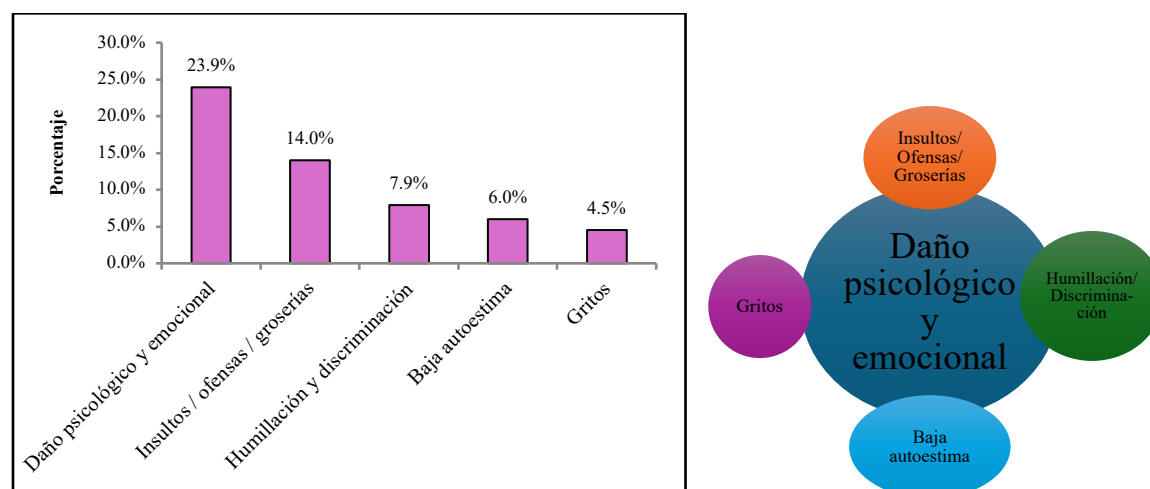
Al analizar el núcleo central de la representación social sobre la violencia verbal se aprecia que predominan los “insultos, ofender y groserías”, y que los elementos periféricos que se derivan del proceso de clasificación son “gritos; agresión verbal; las expresiones tonto(a), pendejo(a), puto(a); y humillación, discriminación y denigración”. Este conjunto de palabras da cuenta de las actuales formas de relacionarse y comunicarse entre los jóvenes universitarios; son manifestaciones que los daña, agrede y ofende. “Las palabras y las emociones nos envuelven, nos tocan, nos envuelven nos incluyen, nos someten y nos marginan (Kaplan, 2016, p. 212)... “Los insultos, las estigmatizaciones racistas, etc., como categoremas, como decía Aristóteles, es decir, acusaciones públicas, actos de designación, de nominación, que pretenden la universalidad, por lo tanto, la autoridad sobre el mundo social (Bourdieu, 1987, p.99)” (Kaplan, 2016, p. 213).

Para afrontar lo anterior, se deben modificar las prácticas sociales perjudiciales, erradicar las expresiones verbales violentas e inculcar en los estudiantes los valores del respeto y la tolerancia. Cabe señalar que la UV forma profesionales con un alto dominio en conocimientos especializados, en determinado campo disciplinario; se les prepara para desempeñarse en diferentes áreas de trabajo. Por ello, se debe fomentar, entre otras cosas, que los estudiantes hablen con propiedad en los espacios universitarios, focalizando sus expresiones en una forma de comunicación armoniosa con los demás.

Violencia psicológica o emocional

Respecto a la violencia psicológica o emocional, el 23.9% señaló el “daño psicológico y emocional”; el 14.0%, los “insultos, ofensas y groserías”; el 7.9%, la “humillación y discriminación”; el 6.0%, la “baja autoestima”; y el 4.5%, “los gritos”.

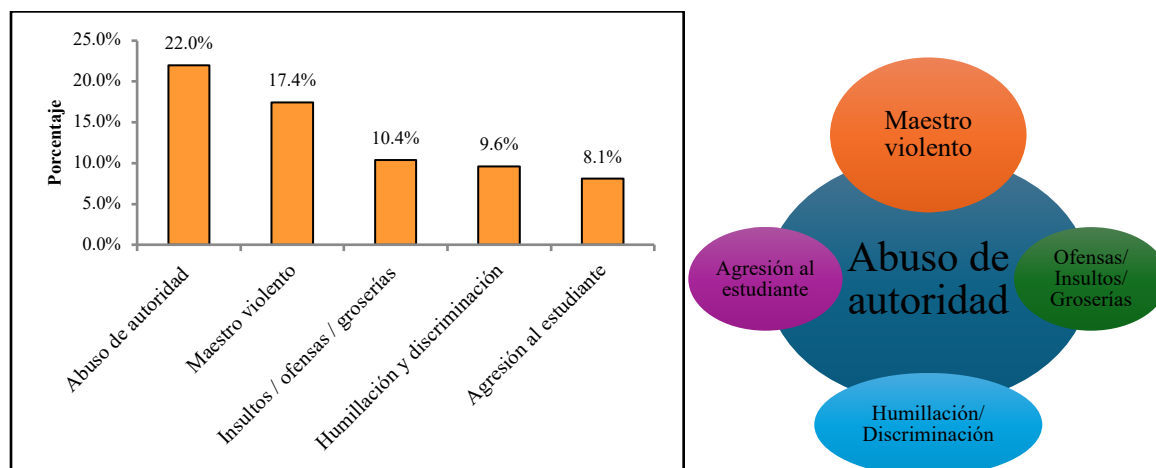
Gráfica 4. Principales palabras que asocian cuando se habla de violencia psicológica o emocional



Se valora que el núcleo central de la violencia psicológica o emocional es “el daño psicológico o emocional”; y los elementos periféricos son “insultos, ofensas, groserías; humillación y discriminación; baja autoestima; y gritos”. Desafortunadamente, se percibe que la presencia del abuso psíquico afecta la autoestima y el desarrollo operativo de los estudiantes, tanto dentro como fuera de la universidad.

Violencia docente

Gráfica 5. Principales palabras que asocian cuando se habla de violencia docente



En lo que atañe a este tipo de violencia, el 22.0% de los estudiantes universitarios expresaron el “abuso de autoridad”; el 17.4% señalaron al “maestro violento”; el 10.4%, “insultos, ofensas y groserías”; el 9.6%, “humillación y discriminación”; y el 8.1% lo asocian con la “agresión al estudiante”. Evidentemente, dos de cada diez estudiantes perciben abuso de autoridad por parte de los docentes en la universidad.

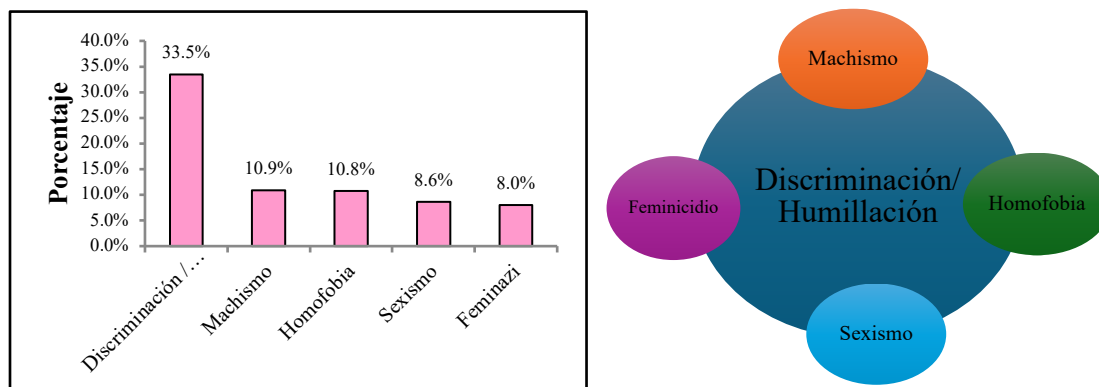
Los testimonios de los universitarios develan que el núcleo central de esta representación es el “abuso de autoridad”, seguido de la categorización de sus elementos periféricos: “maestro violento; insultos, ofensas y groserías; humillación y discriminación; y agresión al estudiante”.

Desgraciadamente, existen algunos profesores dentro de la universidad que ejercen relaciones de poder y control hacia los estudiantes, ya sea por cuestiones de edad, género, condición social, económicas, etc. Lo importante es enseñarles a los estudiantes a identificar el acto violento, directo o tácito, con el objeto de que denuncien los hechos y evitar estas formas de dominación en las aulas universitarias.

Violencia de género

Al plantear la pregunta ¿Cuáles son las cinco palabras que asocias cuando hablamos de violencia de género?, el 33.5% de los estudiantes estimó que es “discriminación, humillación y denigración”; el 10.9% dijo “machismo”; el 10.8%, “homofobia”; el 8.6%, “sexismo”; y el 8.0%, “feminazi”.

Gráfica 6. Principales palabras que asocian cuando se habla de violencia de género



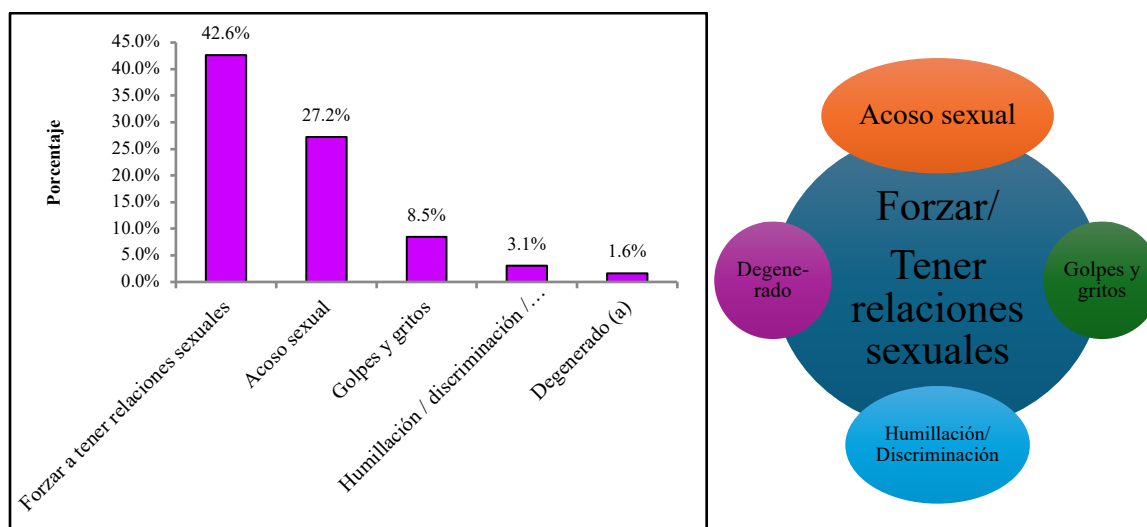
En lo que concierne al núcleo central de la representación social sobre la violencia de género, es la “discriminación”, donde el ordenamiento de sus elementos periféricos es el siguiente: “machismo, homofobia, sexismo y feminazi”.

Lamentablemente, esta violencia afecta a las mujeres por el hecho de serlo, constituyendo un atentado contra su integridad, dignidad y libertad. La agresión se manifiesta en las desigualdades y en el marco de dominación de los hombres sobre las mujeres; son prácticas nocivas en las relaciones humanas que se gestan en la sociedad y que han perdurado en el devenir histórico, donde la universidad no se ve excluida de ellas, dado que muchas mujeres estudiantes llegan a ser subordinadas por el hombre, en ocasiones son sus compañeros, maestros, amigos, novio o pareja quienes las violentan, con el objeto de mostrar su poder sobre de ellas. Particularmente Dorantes (2019), precisa que “las prácticas de violencia de género perduran como elementos culturales que están insertos en las prácticas cotidianas y en las formas de pensar de los estudiantes” (p. 130); y agrega “la violencia de género se encuentra presente en la mente de los estudiantes de la Universidad Veracruzana y de la Universidad Veracruzana Intercultural... Hombres y mujeres muestra un grado de naturalización y apropiación de la violencia de género, conciben que es una forma de discriminación y humillación, y que les parece indignante y les enoja (Dorantes, 2019, p. 133), finalmente indica que dicha violencia “se ve manifestada en actos, expresiones y acciones que se sostienen en lógicas de dominación y de relaciones de poder que logran soldarse como naturales y normales Dorantes, 2019, p. 135).

Violencia sexual

En relación con la pregunta ¿Cuáles son las cinco palabras que asocias cuando hablamos de violencia sexual?, el 42.6% consideró que se vincula con “forzar a tener relaciones sexuales”; el 27.2%, con “acoso sexual”; el 8.5%, con “golpes y gritos”; el 3.1%, con “humillación, discriminación y denigrar”; y el 1.6%, con “degenerado(a)”.

Gráfica 7. Principales palabras que asocian cuando se habla de violencia de sexual

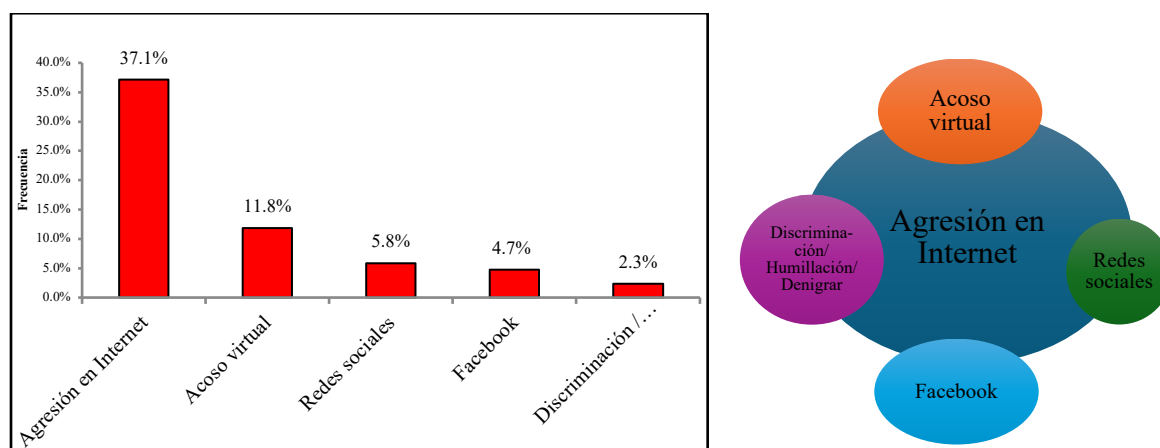


Como se puede observar, el núcleo central de la representación social sobre la violencia sexual es “forzar a tener relaciones sexuales” y la sistematización de sus elementos periféricos es “acoso sexual; golpes y gritos; humillación, discriminación, denigrar; y degenerado(a)”.

Esta violencia remite al daño sexual que se hace a una persona a través de comentarios, insinuaciones no deseadas, acciones para comercializar o utilizar la sexualidad en cualquier ámbito; además, se manifiesta con agresiones por medio de la fuerza física, psíquica o moral. Su objeto es imponer una conducta sexual en contra de la voluntad.

Cyberbullying

En cuanto a la violencia por medio de las redes sociales, el 37.1% de los encuestados afirmó que es “agresión en Internet”; el 11.8%, “acoso virtual”; el 5.8%, “redes sociales”; el 4.7%, “Facebook”; y el 2.3%, “discriminación, humillación y denigrar”.

Gráfica 8. Principales palabras que asocian cuando se habla de *cyberbullying*

Por lo que se refiere al núcleo central de la representación social sobre el *cyberbullying*, se aprecia que predomina “la agresión por Internet” y la categorización de sus elementos periféricos es la siguiente: “acoso virtual; redes sociales; Facebook; discriminación, humillación y denigrar”.

Afortunadamente, los jóvenes tienen mayores oportunidades de utilizar herramientas tecnológicas y accesibilidad para conectarse al Internet; no obstante, los entornos virtuales se han convertido en medios para violentar a los demás. Por tanto, es fundamental cuidar el uso y manejo de las redes sociales, ya que una vez que se causan daños difícilmente se pueden reparar, inclusive la información que se sube a las plataformas digitales puede ser utilizada nuevamente con el tiempo.

CONCLUSIONES

El estudio de las representaciones sociales nos permitió constatar que existe un conjunto de prácticas violentas que ejercen y padecen los estudiantes de la Universidad Veracruzana, de sus diversas regiones y áreas de conocimiento. Los universitarios perciben, experimentan y comparten los ocho tipos de violencia aquí analizados; lo grave es que varias de las manifestaciones violentas se entrelazan y complementan, afectando a los jóvenes de manera simultánea.

En el estudio verificamos que la idea de la centralidad o del núcleo central de la representación social permite comprender y profundizar en la violencia y su tipología. A la vez, los elementos periféricos dan cabida a explicar las causas de la violencia y las formas en cómo se practica este acto. Así, las palabras le otorgan un significado a cada tipo de violencia, a partir de las percepciones y comportamientos de los estudiantes, de su interacción con el entorno y sus experiencias, con lo cual se articula y construye la representación social.

Recapitulando, el núcleo central de las representaciones sociales por cada tipo de violencia es el siguiente:

a) violencia física: “golpes”; b) violencia económica: “falta de dinero”; c) violencia verbal: “insultos, ofender y groserías”; d) violencia psicológica o emocional: “daño psicológico o emocional”; e) violencia docente: “abuso autoridad”; f) violencia de género: “discriminación, humillación y denigración”; g) violencia sexual: “forzar a tener relaciones sexuales”; y h) *cyberbullying*: “agresión en Internet”.

De manera general, se puede decir que el conjunto de representaciones nos permite conocer realidades emergentes. Se pone en escena un sistema de valores predominantemente violento, los cuales reflejan la cultura universitaria y un conjunto de normas que determinan el entorno de la UV, por lo que es necesario atender a la violencia y su tipología de manera focalizada.

Finalmente, el manejo de herramientas estadísticas nos permitió analizar elementos e indicadores con mayor precisión; así mismo, comprender la realidad de manera precisa y cercana, reflexionando sobre las manifestaciones violentas que dañan a los universitarios. El estudio nos invita a reconocer que “existe una necesidad de modificar prácticas, de tener mayor sensibilidad, de atender los múltiples problemas que afectan a los jóvenes universitarios, se debe impulsar la denuncia y cualquier tipo de violencia que sufran tanto hombres como mujeres” (Dorantes, 2017, p. 121).

Podemos concluir que es elemental hacer conciencia de lo que nos perjudica como seres humanos, encontrar mejores formas de relacionarnos unos con otros, aprender a solucionar los conflictos de manera pacífica y tomar decisiones que mejoren el ambiente escolar, con la finalidad de coexistir en una universidad libre de violencia.

REFERENCIAS

- Abric, J. C. (1994). *Prácticas sociales y representaciones*. México: Ed. Coyoacán.
- Álvarez, J. (2004). El contexto social y teórico del surgimiento de la teoría de las representaciones sociales. En Romero, E. (2004). *Representaciones sociales. Atisbos y cavilaciones de cuatro décadas*. México: Benemérita Universidad Autónoma de Puebla (BUAP). (pp. 20-53).
- Arendt, H. (2005). *Sobre la violencia*. Madrid, España: Taurus.
- Banchs, M. A. (2000). *Aproximaciones procesuales y estructurales al estudio de las representaciones sociales*. Venezuela: Universidad Central de Venezuela.
- Bringiotti, M. I., Krynveniuk, M., y Lassi, S. (2005). Las múltiples violencias de la “violencia” en la escuela. Hacia un abordaje integrativo del problema. En G. Averbuj, L. Bozzalla, M. Marina, G.

- Tarantino y G. Zaritzky (Comps.), *Violencia y escuela, propuestas para comprender y actuar* (pp. 25-48). Buenos Aires, Argentina: Aique.
- Buquet, A., Cooper, J., Mingo, A., y Moreno, H. (2013). *Intrusas en la universidad*. México: UNAM, PUEG, IISUE.
- Casillas, M.A; Dorantes, J. J; y Ortiz, V. (2017). *Estudios sobre la violencia de género en la universidad*. Xalapa, Veracruz: Universidad Veracruzana, Biblioteca Digital de Humanidades.
- Casillas, M. A; Dorantes, J. J; Guadarrama, M. A; Castro, C. y Morales, C. (2017). Tipos de violencia en la Universidad: Primeros resultados de una investigación en la UV. En Casillas, M.A; Dorantes, J. J; y Ortiz, V. (2017). *Estudios sobre la violencia de género en la universidad*. Xalapa, Veracruz: Universidad Veracruzana, Biblioteca Digital de Humanidades (pp. 55-77).
- Combessie, J. C. (2003). *El método en sociología*. Argentina: Ferreira.
- Conceptodefinición.de. (2016). *Violencia docente*. Recuperado el 8 de febrero de 2020 en <https://conceptodefinicion.de/violencia-docente/>
- Dorantes, J. J. y Morales, C. (2019). Representaciones sociales de estudiantes de la UVI sobre la violencia de género. En Badillo, J. y Ortiz, V. (2019). (Coordinadoras). *Aproximaciones a la interculturalidad y el género desde la investigación*. México: Red Durango de Investigadores Educativos (pp. 114-140).
- Dorantes, J. J. (2017). La violencia de género, miradas de los estudiantes de la Universidad Veracruzana. En Casillas, M.A; Dorantes, J. J; y Ortiz, V. (2017). *Estudios sobre la violencia de género en la universidad*. Xalapa, Veracruz: Universidad Veracruzana, Biblioteca Digital de Humanidades (pp. 101-123).
- Flament, C. (1994). Estructura, dinámica y transformación de las representaciones sociales. En J. C. Abric (Ed.). *Prácticas sociales y representaciones*. México: Ed. Coyoacán.
- Furlán, A., y Spitzer, T. (2013). Panorama Internacional. En A. Furlán y T. Spitzer (Eds.), *Convivencia, disciplina y violencia en las escuelas* (pp. 39-71). México: ANUIES.
- Jodelet, D., y Guerrero, T. (2000). *Develando la cultura. Estudios en representaciones sociales*. México: UAM.
- Kaplan, C. V. (2016). "El lenguaje es una piel. Género, violencia y procesos civilizatorios." En Kaplan, C. V. (Editora). *Género es más que una palabra. Educar sin etiquetas*. (pp. 211-223). Buenos Aires: Miño y Dávila.

- Marradi, A., Archiento, N., y Piovani, J. I. (2007). *Metodologías de las Ciencias Sociales*. Argentina: Planeta.
- Moscovici, S. (1961). *El psicoanálisis, su imagen y su público*. Buenos Aires, Argentina: Huemul.
- Organización Mundial de la Salud [OMS]. (2020a). *Violencia*. Recuperado el 8 de febrero de 2020 en <https://www.who.int/topics/violence/es/>
- Organización Mundial de la Salud [OMS]. (2020b). *Violencia juvenil*. Recuperado el 9 de febrero de 2020 en <https://www.who.int/es/news-room/fact-sheets/detail/youth-violence>
- Significados.com. (2019). *Tipos de violencia*. Recuperado el 8 de febrero de 2020 en <https://www.significados.com/tipos-de-violencia/>
- Taylor, S. L., y Bogdan, R. (2006). *INTRODUCCIÓN a los métodos cualitativos de investigación*. Barcelona, España: Paidós.
- Villarroel, G. (2007). Las representaciones sociales: una nueva relación entre el individuo y la sociedad. *Fermentum. Revista Venezolana de Sociología y Antropología*, 17(49), 434-454. Recuperado de <https://www.redalyc.org/pdf/705/70504911.pdf>

ANÁLISIS ESTADÍSTICO DE LA RELACIÓN ENTRE LA SINTOMATOLOGÍA DE ANSIEDAD Y DEPRESIÓN CON EL RENDIMIENTO ESCOLAR EN UNA MUESTRA DE ALUMNOS BACHILLERES

Guadalupe Sánchez Sánchez¹, María de Lourdes Velasco Vázquez¹, Blandina Bernal Morales²

¹Facultad de Estadística e Informática, ²Instituto de Neuroetología,

Universidad Veracruzana.

E-mail: ssanchezdel12@gmail.com.

RESUMEN

Objetivo: Identificar si la presencia de sintomatología de ansiedad y depresión se relaciona con las calificaciones en estudiantes de primer año del Centro de Bachillerato Tecnológico industrial y de servicios No. 13 "Dr. José María Cos y Pérez. Desarrollo: El grupo de los denominados trastornos afectivos surge de la convergencia, ocurrida al inicio del siglo XX, de determinados términos y conductas. La palabra afectivo forma parte de una colección de vocablos, como emoción, pasión, sentimiento, ánimo, afecto, distimia, ciclotimia o disforia, que denominan experiencias subjetivas superpuestas, con diferentes etimología y origen (Luque y Berrios, 2011). En diversos estudios se ha reportado que los trastornos mentales constituyen uno de los mayores problemas de salud en todo el mundo (Berenzon y Juárez, 2005). Existen diferentes trastornos del afecto, también conocidos como trastornos del estado de ánimo, entre ellos, el trastorno depresivo y el trastorno de ansiedad (Baena *et al.*, 2005). Los estudiantes que ingresan al bachillerato con los cambios fisiológicos, anatómicos y psicológicos de su edad, representan una población sensible al ajuste que demanda una educación en un nivel de exigencia más avanzado, volviéndolos una población vulnerable a los trastornos afectivos (Barbero, 2002). Conclusión: Existe correlación positiva entre las variables latentes ansiedad y depresión; situación contraria se presenta en las calificaciones y estos trastornos, donde la correlación es negativa. La sintomatología de la depresión se correlaciona con un 18% a las bajas calificaciones, mientras que la sintomatología de ansiedad en un 3%. **Palabras clave:** Depresión. Ansiedad. Trastorno afectivo. Educación Media Superior. Rendimiento Académico.

INTRODUCCIÓN

Los estudiantes que ingresan al bachillerato con los cambios fisiológicos, anatómicos y psicológicos de su edad, representan una población sensible al ajuste que demanda una educación en un nivel de exigencia más avanzado (Barbero, 2002). Ellos viven la adolescencia como una etapa de cambios significativos que los hace vulnerables entre otras cosas, a trastornos de depresión y ansiedad de origen multifactorial (estatus económico familiar bajo, problemática ambiental tanto social como geográfica, los problemas de salud, etc.) que pueden repercutir en su rendimiento escolar. El problema es que en la población adolescente, la depresión o la ansiedad frecuentemente es subdiagnosticada, ignorada, o no detectada, porque puede coexistir con otros trastornos (trastornos de alimentación, abuso de drogas, conducta antisocial) o simplemente ser sujeto de indiferencia en las personas que los rodean (Arenas Ayllon, 2009; Hoyos *et al.*, 2012). Se ha documentado que la depresión y ansiedad en adolescentes se asocia a una mayor probabilidad de conducta violenta, así como al bajo rendimiento escolar y al consumo de sustancias adictivas, lo que pone en riesgo la continuidad de su vida personal y profesional futura. Por ello es que este trabajo pretende explorar la sintomatología de depresión y ansiedad en una población adolescente en un bachillerato local y busca una posible relación con algunos aspectos de rendimiento escolar (Aguilar Sánchez, 2014). Este trabajo consistirá en un estudio exploratorio de la sintomatología emocional y de la relación estadística con las variable escolar calificaciones de las materias durante el semestre, en estudiantes de primer año del Centro de Bachillerato Tecnológico industrial y de servicios No. 13 "Dr. José María Cos y Pérez", en la ciudad de Xalapa, Veracruz, para contribuir con un mejor conocimiento de la población adolescente que ha sido escasamente atendida.

MÉTODO

Se realizó un estudio de encuesta, de corte cuantitativo con alumnos de la generación 2018- 2021 de una escuela de nivel medio superior. Se hizo un muestreo por conveniencia de 222 alumnos (pertenecientes al turno vespertino). Se aplicaron las dos escalas impresas para medir depresión y ansiedad. El llenado del cuestionario se realizó de manera individual y la información obtenida es de carácter confidencial. De los alumnos seleccionados en el muestreo se elaboró un compendio de las calificaciones de algunas de las materias llevadas durante el semestre. Los datos se analizaron mediante un Análisis factorial Confirmatorio para determinar la relación entre las variables.

Población de estudio

La población de estudio fue de 222 estudiantes del primer año de todas las áreas (Mecánica Industrial, Administración de Recursos Humanos, Contabilidad, Logística y Laboratorista Clínico) del Centro de Bachillerato Tecnológico industrial y de servicios No. 13 "Dr. José María Cos y Pérez" turno vespertino en la ciudad de Xalapa, Ver., con una edad entre 15 y 18 años.

Criterios de inclusión:

- 1.- Individuos de entre 15 y 17 años.
- 2.-Individuos cuyos padres firmen la carta de consentimiento informado.
- 5.- Deben hablar español, saber leer y escribir, estar sanos o con sintomatología depresiva o ansiosa.

Criterios de exclusión:

- 1.- Individuos mayores de 18 años.

Instrumentos de medición Escala de Zung

Cuestionario para medir sintomatología de depresión, consta de 20 ítems con escala de Likert, validada en castellano y utilizada en estudios con población mexicana (Conde et al., 1970, Pulido-Criollo y Rodríguez-Landa, 2006) con una alfa de Cronbach de 0.770 (Hernández- Pozo, 2008). El nivel de severidad de los síntomas puede ser No depresión (< 28 puntos), Depresión leve (28-41 puntos), Depresión moderada (42-53 puntos), Depresión grave (> 53 puntos). En esta escala se pueden analizar si se desea los resultados en dimensiones clasificadas como los síntomas somáticos, los síntomas cognoscitivos, síntomas del estado del ánimo y de ansiedad, retardo psicomotor y baja autoestima y disminución del apetito (Cogollo et al., 2006; Hernández-Pozo et al., 2008).

Escala BSPS de ansiedad o fobia social

Cuestionario para hacer una exploración preliminar de ansiedad, formado por 17 ítems con opciones de respuesta en escala de Likert. El punto de corte propuesto para la ausencia de ansiedad es 18. Puntuaciones mayores de 20 se corresponden a un paciente con fobia social que requiere tratamiento (Davidson et al., 1991). La escala constituye un instrumento de evaluación sencillo y rápido de aplicar. Se trata de un cuestionario en formato de escala que mide los síntomas. Los ítems se contestan en una escala de 5 puntos (entre 0 y 4). Los primeros 13 ítems aluden a situaciones sociales específicas, formado por dos constructos que evalúan dos componentes: Miedo (6 ítems: 1,3,5,10,14,15) y Evitación ante las situaciones sociales

(7 ítems: 4,6,8,9,11,12,16). Los cuatro ítems (2,7,13,17) restantes miden respuestas psicofisiológicas. Fundamentalmente, se indaga sobre la posible presencia de enrojecimiento facial, palpitaciones, temblores y sudoración en el transcurso de situaciones interpersonales o cuando se piensa en ellas. Se puede obtener una puntuación total, así como tres puntuaciones relativas a las tres subescalas: temor, evitación y fisiológica (Zubeidat *et al.*, 2006; Cáez-Reyes *et al.*, 2013).

Procesamiento estadístico

Los instrumentos psicológicos fueron validados de forma cuantitativa a través de la consistencia interna con el Alfa de Cronbach con la librería psych del software libre R versión

3.5.1. Los datos obtenidos de los cuestionarios de ansiedad y depresión y los resultados de las calificaciones del primer semestre fueron sometidos a un análisis estadístico descriptivo. También se realizó un análisis inferencial de los puntajes de los cuestionarios de ansiedad y de depresión, que nos permita hacer predicciones acerca de la población y la sintomatología asociada a las calificaciones. Se aplicó una técnica multivariante que permite estimar el efecto y las relaciones entre variables latentes; depresión, ansiedad y calificaciones de una manera simultánea, con la posibilidad de analizar los constructos y parcelas de los cuestionarios con las calificaciones de geometría, inglés, química y lectura, empleando el análisis factorial confirmatorio utilizando librería Lavaan del software libre R versión 3.5.1.

RESULTADO

Tabla 1. Índices de ajuste para los modelos estructurales

Modelo	X ²	<i>p</i>	CFI	RMSEA	AIC	wAIC
Inicial	73.09	0.002	0.97	0.06	5217.0	0.012
Propuesto	60.33	0.016	0.98	0.05	5208.2	0.987

Nota: CFI= Índice de ajuste comparativo (>0.95); RMSEA=Erro cuadrático medio de aproximación (<0.06); AIC= Criterio de Información de Akaike; wAIC= Criterio de información de Akaike consistente.

Se elaboró un estudio de contraste entre el modelo inicial (gráfico 1) y un modelo propuesto (gráfico 2), este último incluye la correlación entre las varianzas específicas de las asignaturas inglés y lectura, también señala que existe una carga factorial entre la depresión y dos indicadores de la ansiedad, por otro lado se propone que la correlación entre ansiedad y calificaciones no presenta relevancia notoria, como el caso de la variable latente depresión y calificaciones.

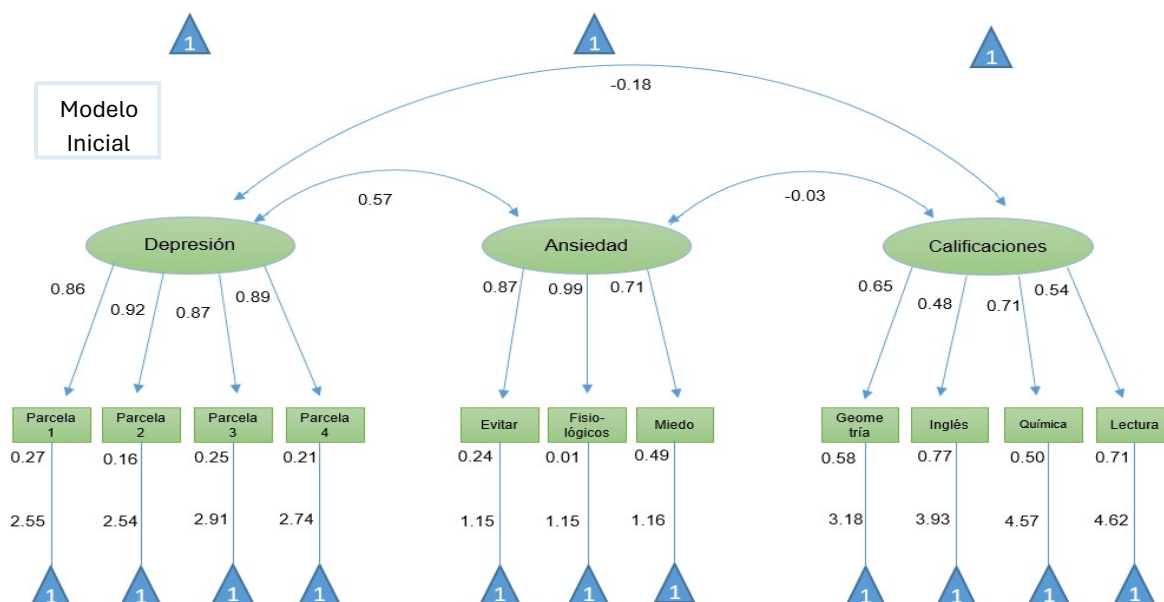


Gráfico 1. El modelo estandarizado inicial indicó un efecto negativo de la sintomatología de ansiedad y depresión sobre las calificaciones, con una correlación= -0.03 y -0.18, respectivamente. La variable depresión explica mejor la parcela 2 = 0.92, para ansiedad es el indicador Evitar= 0.99 y en las calificaciones la materia de Química= 0.71.

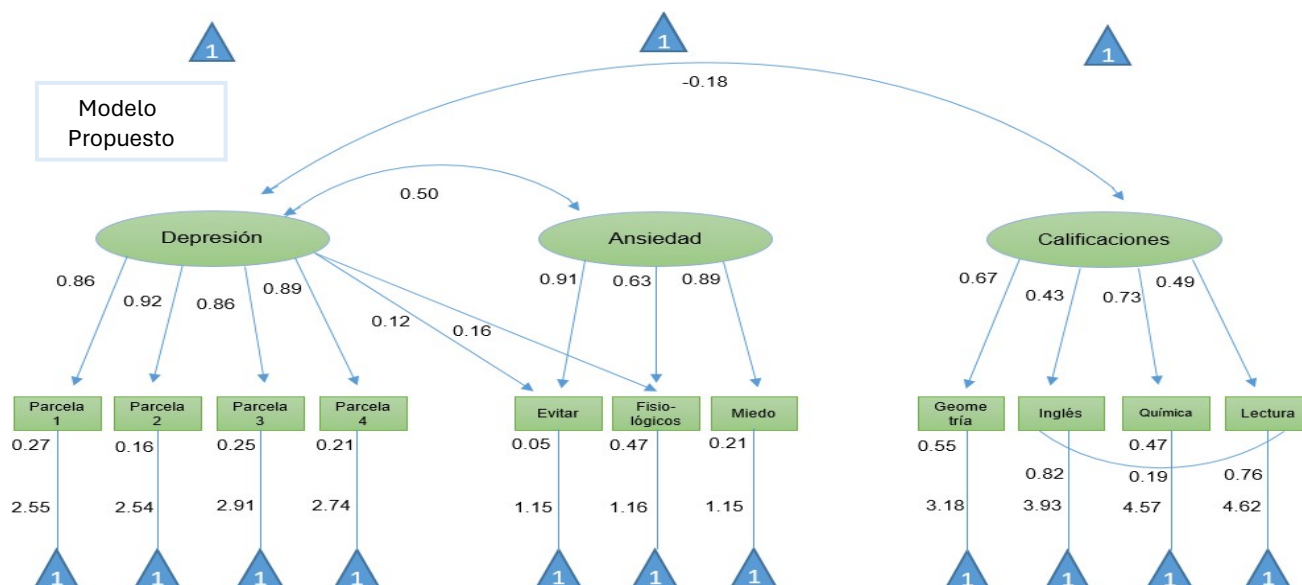


Gráfico 2. El modelo estandarizado propuesto indicó que la sintomatología de la depresión podría compartir los indicadores de Evitar y Fisiológicos de la ansiedad. En la variable Calificaciones, las materias de inglés y lectura presentan una correlación positiva.

CONCLUSIÓN

La presencia de sintomatología de ansiedad y depresión se relaciona con las calificaciones en estudiantes de primer año del Centro de Bachillerato Tecnológico industrial y de servicios No. 13 "Dr. José María Cos y Pérez.

REFERENCIAS

- Aguilar Sánchez, M.L. (2014). *Trastorno depresivo y suicidio* (Tesis doctoral). Departamento de Psiquiatría y Psicología Médica. Universidad de Salamanca. Salamanca, España.
- Arenas Ayllón, S.C. (2009). *Relación entre la funcionalidad familiar y la depresión en adolescentes* (Tesis de Licenciatura). Universidad Nacional Mayor de San Marcos Facultad De Psicología. Lima, Perú.
- Baena Zúñiga, A., Sandoval Villegas, M.A., Urbina Torres, C.C., Helen Juárez, N., Villaseñor Bayardo, S.J. (2005). Los trastornos del Estado de Ánimo. *Revista Digital Universitaria*, 6(11), 2-14. Recuperado de http://www.revista.unam.mx/vol.6/num11/art110/nov_art110.pdf
- Barbero, J.M. (2002). Jóvenes: comunicación e identidad. *Pensar Iberoamérica Revista de Cultura*. Recuperado de <http://www.oei.es/historico/pensariberoamerica/ric00a03.htm>
- Berenzon, S. & Juárez, F. (2005). Mental health services to por patients whith affective disorders in Mexico City, *Revista Saúde Pública*; 39 (4):619-26.
- Cáez-Reyes, K.L., Sayers-Montalvo, S.K., Martínez-Taboas, A. (2013). Validación del Inventario de Fobia Social en una muestra de personas con trastorno de ansiedad social en Puerto Rico. *Revista Puertorriqueña de Psicología*; 24(2).
- Cárdenas, E.M., Feria, M., Vázquez, J., Palacios L., De la Peña, F. (2010). *Guía Clínica para los Trastornos Afectivos (Depresión y Distimia) en Niños y Adolescentes*. Ed. S Berenzon, J Del Bosque, J Alfaro, ME Medina-Mora. México: Instituto Nacional de Psiquiatría. (Serie: Guías Clínicas para la Atención de Trastornos Mentales).
- Cogollo, Z., Díaz, CE., Campo, A. (2006). Exploración de la validez de constructo de la escala de Zung para depresión en adolescentes escolarizados. *Colombia Médica*, 37, 102- 106.
- Conde, V., Escribá, J.A., Izquierdo, J. (1970). Evaluación estadística y adaptación castellana de la escala autoaplicada para la depresión de Zung. *Archivos de Neurobiología*, 33: 185- 206.
- Davidson, J.R.T., Potts, N.L.S., Richichi, E.A., Ford, S.M., Krishnan, K.R.R., Smith, R.D. & Wilson, W. (1991). The Brief Social Phobia Scale. *Journal of Clinical Psychiatry*, 52,48- 51.

- Hernández-Pozo, M., Macías, D., Calleja, N., Cerezo, S., del Valle Chauvet, C. (2008). Propiedades psicométricas del inventario Zung del estado de ansiedad con mexicanos. *Psychologia. Avances de la disciplina*, 2(2), 19-46. Recuperado de <http://www.redalyc.org/articulo.oa?id=297225162009>> ISSN 1900-2386.
- Hoyos, E., Lemos, M., Torres, Y. (2012). Factores de Riesgo y de Protección de la Depresión en los Adolescentes de la Ciudad de Medellín. *International Journal of Psychological Research*, 5(1), 109-121.
- Luque, R., Berrios, G.E., (2011). Historia de los trastornos afectivos. *Revista Colombiana de Psiquiatría*, 40(5), 130-146.
- Pulido-Criollo, F., Rodríguez-Landa, J.F. (2006). *Impacto de un taller de alfarería sobre la depresión de mujeres reclusas: Escala de depresión de Zung*. Libro conmemorativo del II Congreso Regional de la Sociedad Interamericana de Psicología. La Habana: Sociedad Interamericana de Psicología: 101-110.
- Zubeidat, I., Fernández Parra, A., Sierra, J. (2006). Ansiedad y Fobia Social: Revisión de los Autoinformes más Utilizados en Población Adulta e Infanto-Juvenil. *Terapia Psicológica*, 24 (1), 71-86.

EVALUACIÓN DE LA MICROPROPAGACIÓN IN VITRO DE HÍBRIDOS SELECTOS DEL GÉNERO VANILLA USANDO UN DISEÑO COMPLETAMENTE AL AZAR

Sulayka Florencia Castelán Culebro^{1, 2}, Jesús Hernández Suárez¹, Rebeca Menchaca García²,
Noé Velázquez Rosas².

¹Especialización en Métodos Estadísticos, ² Centro de Investigaciones Tropicales.

Universidad Veracruzana.

castelansulayka@gmail.com

INTRODUCCIÓN

La Vainilla (*Vanilla planifolia*) es una de las especies más importantes a nivel comercial, utilizada en diversas industrias como la alimentaria, refresquera, etc.; por ello, alcanza altos valores en el mercado internacional. El Cultivo mundial de esta especie es realizada en su mayoría por prácticas de propagación asexual, donde realizan cortes de tallos llamados “esquejes” y son sembrados en tutores. Actualmente la vainilla presenta una alta erosión genética, lo que significa que estas plantas están propensas a la pérdida de grandes cultivares debido a la escasa capacidad de responder ante distintas condiciones ambientales, enfermedades, plagas, etc. (Soto, 2006; Menchaca, 2010).

Debido a los problemas que enfrentan actualmente las plantaciones de vainilla, es importante desarrollar un programa de mejoramiento genético, en donde se evalúen, híbridos de especies silvestres de nuestro país, apoyando al rescate de variedades de vainilla y especies silvestres.

Por ello es importante desarrollar programas de conservación, que rescaten material genético, ya sea de cultivares de diferentes regiones o poblaciones silvestres, como también otras especies aromáticas nativas del género *Vanilla*, como lo es *V. pompona*, para la creación de bancos de germoplasma, los cuales se encargan de resguardar el material genético. De esta manera se pueden seleccionar especies que aporten características deseables para el cultivo (Imagen 1) y realizar recombinación genética para la creación de híbridos (Menchaca, 2010).

La hibridación de especies del género *Vanilla* es dirigida a la búsqueda de caracteres seleccionados a través de la recombinación genética, que permitan la incorporación de rasgos deseables para la especie, como porcentajes más altos de polinización, resistencia a plagas y enfermedad, tolerancia a sequía, mayor crecimiento para generar esquejes en menor tiempo para su propagación clonal (Gutiérrez et al., 2003;

Vallejo y Estrada, 2002). Una vez generados los híbridos, es necesaria la propagación clonal de cada una de las líneas obtenidas, para su resguardo en bancos de germoplasma y su evaluación *in situ* y *ex situ*, lo que creando líneas genéticas mejoradas.

Por lo que, en el presente trabajo, se pretende realizar una evaluación de la reproducción *in vitro* a través del análisis de un diseño experimental completamente al azar, para comparar estadísticamente el crecimiento de dos híbridos selectos del género *Vanilla* con sus parentales, a través de la propagación clonal de segmentos nodales, utilizando un diseño de factorial de tratamientos.

HIPÓTESIS

En condiciones de cultivo *in vitro*, se espera un mayor crecimiento en el híbrido resultado de la línea parental de *Vanilla planifolia* (gameto masculino) y *Vanilla pompona* (gameto femenino), ya que morfológicamente predominan las características maternas. Los medios de cultivo suplementados con Auxinas van a tener un mayor efecto en el crecimiento de yemas axilares de *Vanilla*, en comparación con los otros medios de cultivo.

OBJETIVO GENERAL

-Evaluar el crecimiento inicial de híbridos selectos de vainilla como estrategia de mejoramiento genético de una especie de importancia comercial.

OBJETIVO ESPECIFICO

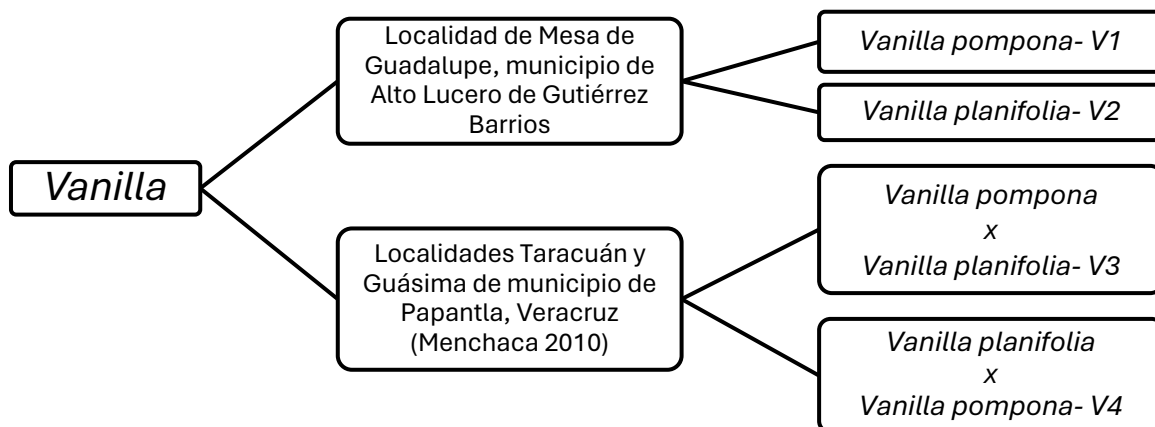
- Seleccionar híbridos con características de mayor crecimiento y resistencia a estrés hídrico.
- Determinar la influencia de los factores auxina, citocinina y plantas del género *Vanilla* en el crecimiento de segmentos nodales.
- Determinar cuál combinación de los factores y niveles optimiza el crecimiento de segmentos nodales.
- Estimar el efecto de la interacción entre los tres factores, auxina, citocinina y plantas del género *Vanilla*
- Determinar la asociación entre el factor de interés (crecimiento) entre la oxidación y contaminación

METODOLOGÍA

La materia vegetal a utilizar serán explantes de segmentos nodales de plantas pertenecientes al género *Vanilla* (Mapa 1). Las cruas entre *V. planifolia* y *V. pompona* han reportado que son compatibles y producen semillas viables para su germinación *in vitro*, así como también muestran un mayor vigor en

crecimiento y la posibilidad de generar esquejes de manera clonal en menor tiempo, en comparación con las plantas que se obtiene por autopolinización, (Menchaca, 2010).

Las cuatro plantas de estudio se encuentran en el banco de germoplasma del Orquidario UV.



Mapa 1.-Procedencia de plantas del género *Vanilla* para extracción de ápices.

Se utilizaron secciones de esquejes para el establecimiento de cultivos axénicos *in vitro*, los cuales van a ser seleccionados observando que no muestren ninguna señal de plaga o enfermedad. Una vez obtenidos, en el laboratorio se cortaron trozos de tallo de aproximadamente 2 cm de largo, cubriendo la zona de la yema, para extraer el ápice y sembrarlos a partir de técnicas de cultivo *in vitro* (Imagen 2).

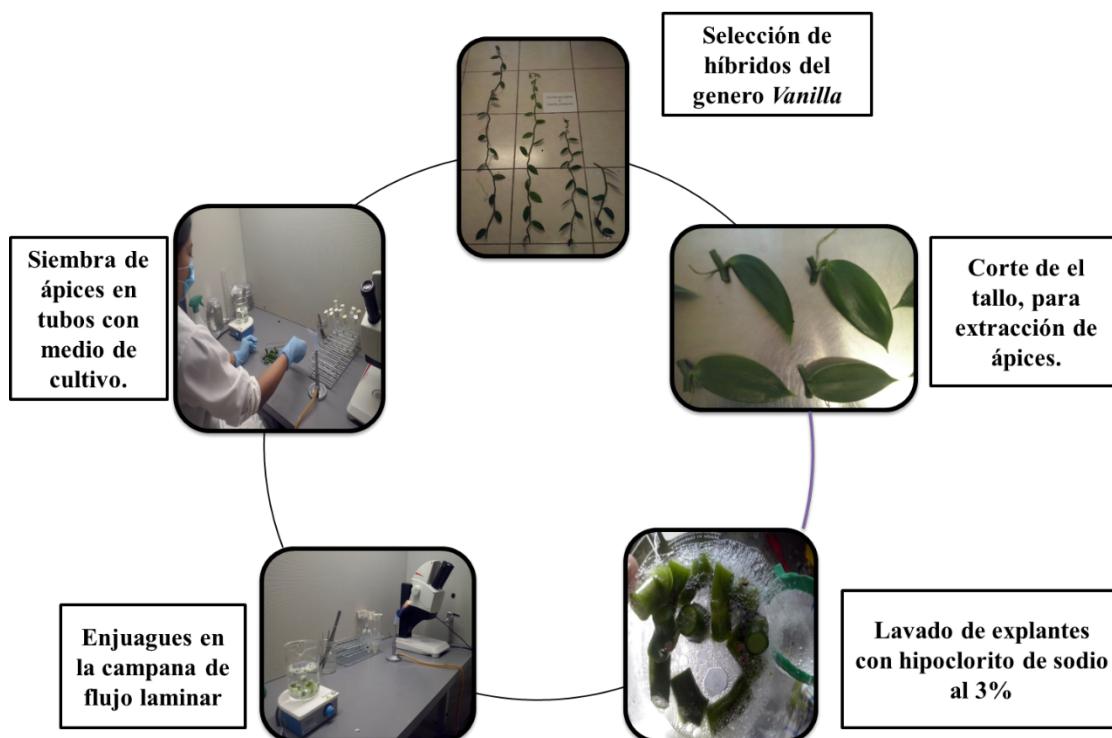


Imagen 2. – Proceso de cultivo *in vitro* de ápices de plantas del genero *Vanilla*

Diseño de Tratamientos

Para estudiar el factor de interés (crecimiento) se va a utilizar un diseño factorial de tratamientos (Kuehl, 2001):

$$P \cdot Q \cdot L$$

Donde:

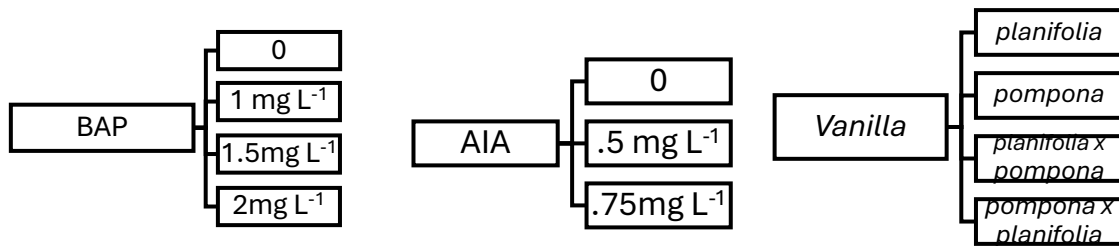
P: =Concentración (mg L^{-1}) de BAP

Q: = Concentración (mg L^{-1}) de AIA

L: = Plantas del genero *Vanilla*

Con un total de 48 tratamientos y 5 réplicas por tratamiento.

Para el establecimiento de yemas *in vitro* se va a utilizar como base el medio de cultivo Murashige y Skoog (MS) con vitaminas (tiamina 0.25 mg, piridoxina 0.25 mg, ácido nicotínico 0.25 mg) adicionado con sacarosa 30g L^{-1} y suplementado con distintas concentraciones de citocininas y auxinas como se muestra en el cuadro 2. El pH va ser ajustando a 5.6 ± 0.01 con hidróxido de sodio (NaOH) y/o ácido clorhídrico (HCl) 1N y solidificado con 3g L^{-1} de gelrite



Cuadro 2. Concentraciones de reguladores de crecimiento: 6- Bencilaminopurina (BAP) y Ácido indolacético (AIA)

Los cultivos se incubarán a una temperatura de $25 \pm 2^{\circ}\text{C}$, con un fotoperiodo de 16 horas de luz y 8 horas de oscuridad en el cuarto de incubación del laboratorio del orquidario UV.

Diseño experimental

Para una comparación precisa en los tratamientos y reducir el error experimental de estudio, se va a utilizar un diseño experimental completamente al azar, donde los ápices extraídos de las zonas maristematicas de la yema axilar son las unidades experimentales, con un total de 240, los cuales fueron colocados de manera aleatoria en las gradillas, con el fin de homogenizar el experimento y reducir el error experimental (ϵ). La toma de datos se realizó cada 10 días.

Las variables a que van a ser evaluadas son; el índice de crecimiento, oxidación y contaminación.

Tasa o índice de crecimiento

La tasa de crecimiento relativo es el el incremento de biomasa por unidad de biomasa y tiempo.

$$TCR = \frac{M_f - M_i}{T_f - T_i}$$

Donde:

Mf=Medición final

Mi=Medición inicial

Tf= Tiempo final

Ti=Tiempo final

Análisis estadístico

Se probarán las siguientes hipótesis:

H ₀ : P=0	Vs	H _a : P≠0
H ₀ : Q=0	Vs	H _a : Q≠0
H ₀ : L=0	Vs	H _a : L≠0
H ₀ : PQ=0	Vs	H _a : PQ≠0
H ₀ : PL=0	Vs	H _a : PL≠0
H ₀ : QL=0	Vs	H _a : QL≠0
H ₀ : PQL=0	Vs	H _a : PQL≠0

Para contrastar las hipótesis se utilizarán un análisis de varianza, en el caso de que se rechace una de las hipótesis, se utilizará la prueba de Dunnet para comparar a todos los tratamientos contra el control.

Modelo estadístico

$$y = \mu + P + Q + L + PQ + PL + QL + PQL + \varepsilon$$

y = tasa de crecimiento relativo

μ = Media general

P: = Concentración (mg L⁻¹) de BAP

Q: = Concentración (mg L⁻¹) de AIA

L: = Plantas del genero *Vanilla*

ε : Error aleatorio

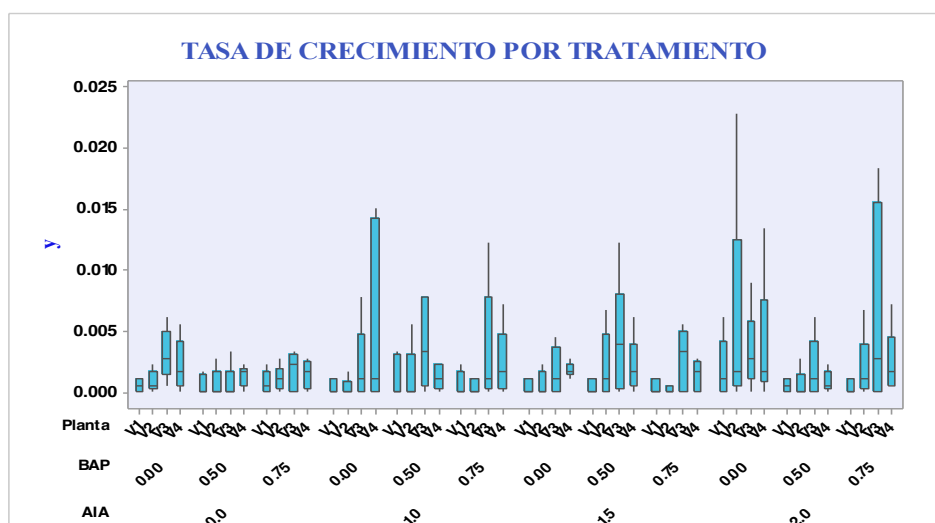
Donde:

$$\varepsilon \sim NID(0, \sigma^2)$$

En caso de que se rechace una H₀, se utilizará una prueba de ordenamiento de medias y se verificaron los supuestos de normalidad, independencia y homocedasticidad, a través de técnicas no analíticas y si procede analíticas. Se analizaron las variables de oxidación y contaminación a través de modelos log-lineales (Graham, 2017)

RESULTADOS

El tratamiento compuesto por el testigo con la planta V3 (V. pompona x V. planifolia), obtuvo el mayor crecimiento, seguido de V3 con 0.75 de AIA, en lo contrario, los tratamientos V1 (testigo) son los que tuvieron menor tasa de crecimiento (Gráfica 1)

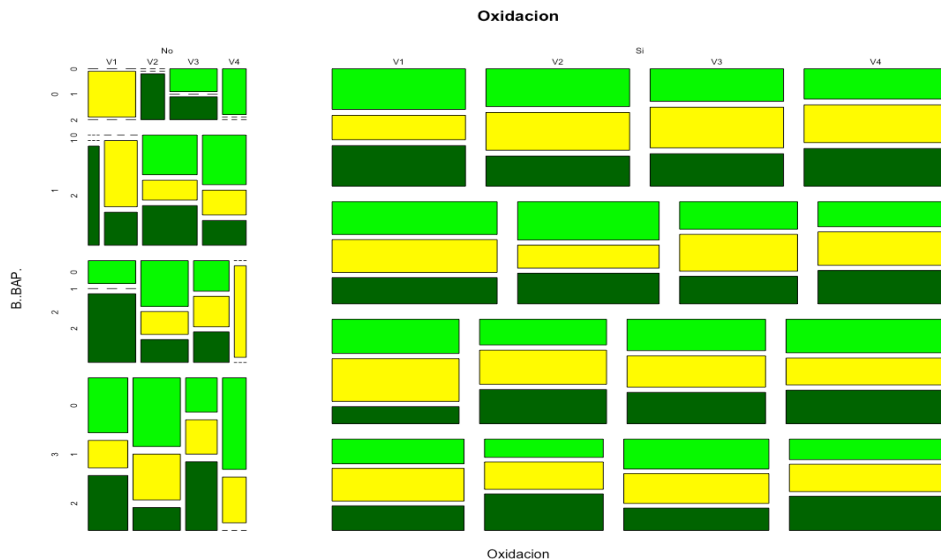
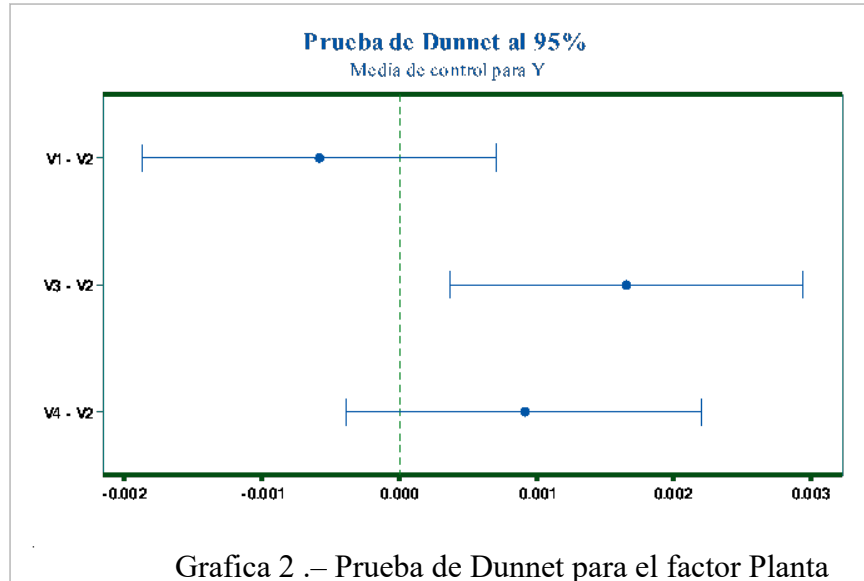


Gráfica 1 .– Comparación de tratamientos en la tasa de crecimiento

TRANSICIÓN DE ESTADÍSTICA A CIENCIA DE DATOS

Las particiones de sumas de cuadrados y la tabla del análisis de varianza (Anexo 1) para este experimento con tres factores, en ella observamos que el estadístico de prueba F conduce al rechazo de la hipótesis nula para el efecto principal “PLANTA” ($GL=3$, $F= 3.35$, $p\text{-valor}=0.020$).

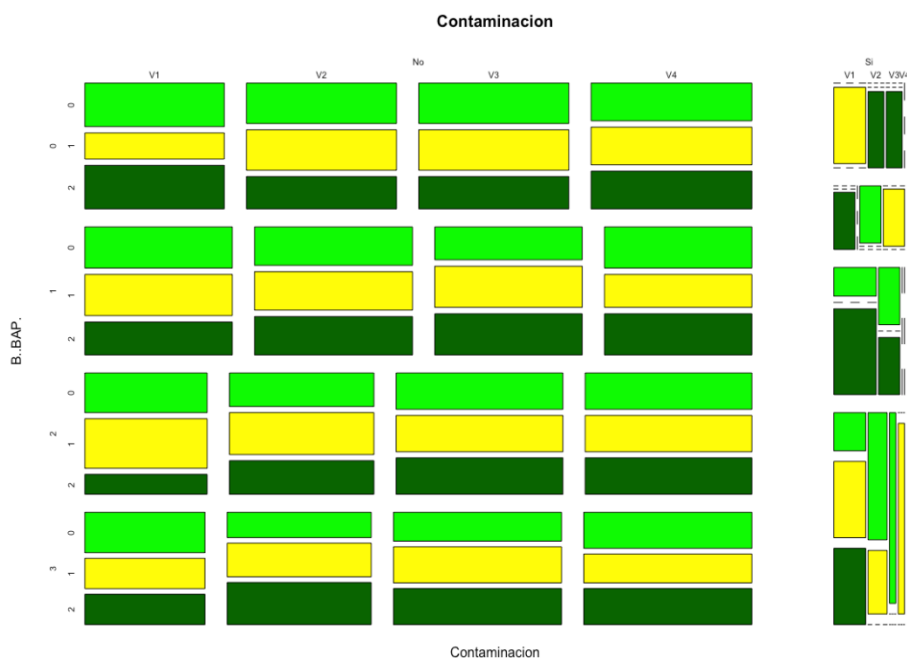
A través de la prueba de Dunnet se compararon los tratamientos, en donde se observa que la planta V3 es diferente de las plantas V1, V2 y V4. Estas últimas plantas son iguales entre ellos estadísticamente (Gráfica 2)



Gráfica 3 .– frecuencia de oxidación por tratamientos

Oxidación

Modelo (BAP*Oxi, Planta*Oxi). Donde BAP es independiente de planta, para cada valor en la oxidación, es decir que si se presenta o no oxidación es independiente de BAP y Planta. ($\chi^2=40.826$, $gl=82$, $p\text{-value}\approx 1$), Grafico 3.



Gráfica 4 .– Frecuencia de contaminación por tratamientos

Contaminación

Modelo (AIA*Contaminación, Planta*Contaminación) AIA es independiente de planta, para cada valor en la oxidación, es decir que si se presenta o no contaminación es independiente de AIA y Planta. ($\chi^2=46.343$, $gl=84$, $p\text{-value}\approx 1$).

CONCLUSIONES

Se acepta la hipótesis propuesta, ya que, a través de los análisis estadísticos, se demuestra que el factor que influye en el crecimiento de las plántulas, es el factor “planta” y no los factores de BAP, AIA. En donde el híbrido producto del resultado de la cruce de *Vanilla pompona* X *Vanilla planifolia* (V3), es diferente de las plantas V1, V2 y V4, como también es el híbrido que obtuvo el mayor crecimiento.

REFERENCIAS

- Menchaca- García, R. A. (2010). Obtención y Caracterización Morfológica de híbridos de Vanilla planifolia G. Jackson in Andrews y V. Pompona Schide. Tesis de doctorado para obtener el título de Doctor en Ecología Tropical. Centro de Investigaciones Tropicales, Universidad Veracruzana, Veracruz, México.
- Gutiérrez, A.M., Santacruz F.R., Cabrera P.J., Rodríguez G. 2003 Mejoramiento genético in vitro e-Gnosis Universidad de Guadalajara, México. 1 (1) 10-12.
- Graham J.G.Uptom. 2017. Categorical data analysis by example, published by John Wiley & Sons. Inc. Hoboken, New Jersey.
- Soto- Arenas, M. A. (2006). La vainilla; retos y perspectivas de su cultivo. CONABIO. Biodiversitas 66: 1-9.
- Kuehl Robert U. 2001. Diseño de Experimentos. Principios estadísticos de diseño o análisis de investigación. Segunda Edición. The University of Arizona.
- Vallejo- cabrera F. A., y Estrada- Salazar E. i. (2002). Mejoramiento genético de plantas. Universidad Nacional de Colombia. Sede Palmira.404 págs.

ANEXO 1

Tabla del análisis de varianza

Fuente de variación	GL	Suma de cuadrados	Cuadrados medios	F	P- valor
AIA	3	0.000070	0.000023	1.31	0.274
BAP	2	0.000074	0.000037	2.08	0.128
PLANTA	3	0.000179	0.000060	3.35	0.020**
AIA*BAP	6	0.000123	0.000020	1.15	0.338
AIA*PLANTA	9	0.000054	0.000020	0.34	0.962
BAP*PLANTA	6	0.000013	0.000002	0.12	0.994
AIA*BAP*PLANTA	18	0.000319	0.000018	0.99	0.468

REPRESENTACIONES SOCIALES DE ACADÉMICOS DE LA UV SOBRE LA VIOLENCIA. NUBE DE PALABRAS Y ANÁLISIS DE SIMILITUD-*SOFTWARE* IRAMUTEQ

Jeysira Jacqueline Dorantes Carrión

Universidad Veracruzana

jeysira@hotmail.com

RESUMEN

En el marco del IV Coloquio Nacional de Investigaciones en Representaciones Sociales, realizado en Xalapa, Veracruz, por el Centro de Estudios de Opinión y Análisis (CEOA) de la Universidad Veracruzana (UV), en el 2015, se impartió un taller para académicos sobre las representaciones sociales (RS), por Ana María Silvana de Rosa, de la Universidad Sapienza de Roma y European PhD on Social Representations & Communications. En este artículo se muestran resultados de una investigación llevada a cabo, a partir de dicho taller, sobre las RS de 21 académicos sobre la violencia o actos violentos “entre escolares, que se detectan en el interior de los centros educativos” (Harris y Petrie, 2006, pp. 11-12); violencia que responde a una estructura, sostenida en una cultura edificada por la lógica de la dominación y por las relaciones de poder (Arisó y Mérida, 2010, p. 10). El instrumento de investigación fue una ficha o tarjeta (De Rosa, 2005) donde los participantes escribieron palabras que relacionaron con el término violencia, las cuales posteriormente fueron clasificadas según el orden de importancia. El análisis fue mediante el *software* Iramuteq (Interfaz de R para el Análisis Multidimensional) (Molina-Neira, 2017); como resultado se obtuvieron gráficos de nubes de palabras y un análisis de similitud. El programa estadístico Iramuteq identificó los núcleos semánticos detectados por coocurrencias de palabras y el núcleo central (Abrieu, 1994) de la RS sobre la violencia. Los resultados indicaron que “golpe” es el núcleo central de la representación social de la violencia y las palabras clave: agresión, maltrato, insulto y daño psicológico son las que se articulan a los elementos periféricos de dicha RS.

Palabras clave: académicos, núcleo central, representaciones sociales, violencia.

INTRODUCCIÓN

Las teorías de la violencia permiten reflexionar que en los asuntos humanos surgen relaciones violentas con inmensas líneas de conductas y situaciones diversas; hacen observables relaciones de poder, que poseen significados del mundo social, del origen del mal y de las desigualdades sociales (Jacorzynski, 2003); asimismo, ayudan a explicar que la violencia es un atributo humano (Sanmartín, 2010). Actualmente, la violencia ha despertado la atención de diversos investigadores, como: Morales, Serrano, Miranda y Santos (2014); Johnson y Johnson (2010); Abramovay (2006); Almeida y Gómez (2005); Novo y Seijo (2009); Grettiez (2008); Furlán (2005); Furlán, Pasillas, Spitzer y Gómez (2010); Blaya (2006); Guillotte (2003); Sofsky (2006); Torres (2005); Rojas-Solís (2013); Arendt (2005); Blair (2009); Castro y Reta (2013), por mencionar a algunos.

Reconocemos que, a nivel nacional e internacional, se han despertado diversas preocupaciones sobre las incidencias y afectaciones de la violencia en hombres y mujeres, pero también en los contextos escolares, incluyendo el universitario. En este último llegan a percibirse distintos rostros y rasgos de la violencia como temáticas centrales que se discuten entre los académicos e investigadores. Surge una preocupación por lo que acontece en los contextos universitarios, y nos preguntamos ¿cuáles son las experiencias de violencia que sufren y experimentan los estudiantes universitarios?, y ¿cuáles son los actos de violencia que más perciben, específicamente en la Universidad Veracruzana (UV)?

Es importante valorar que en la UV existe el *Programa de Trabajo Estratégico 2013-2017 Tradición e Innovación* (Universidad Veracruzana, 2013), cuya prioridad es atender la violencia en la Instituciones de Educación Superior (IES), pero también rechazar los actos violentos y la discriminación. Hoy damos cuenta de una complejidad de las relaciones entre los y las estudiantes, y también entre los profesores y las profesoras, solo que no existen diagnósticos claros que den cuenta de ello, sino que son formas manifiestas de la violencia que se entrelazan en las relaciones humanas y suelen ser destructivas, dañinas, que detonan experiencias muchas veces dolorosas.

Dorantes (2017) plantea que los estudiantes universitarios experimentan, de manera simultánea, la violencia verbal, docente, física, psicológica o emocional, de género, económica, *cyberbullying*, sexual, escolar, administrativa y familiar; inclusive, reconocen que procede de estudiantes, profesores(as), compañeros(as) de grupo, amigos(as), novios(as), personal administrativo, de intendencia y vigilancia, con los que interactúan en los espacios físicos de la universidad. Las edades de los estudiantes que sufren violencia de género son de los 18 a 21 años. Aunque los porcentajes no son muy altos, la violencia tiene distintas manifestaciones y un impacto en los jóvenes universitarios (Dorantes, 2017).

Los resultados del estudio describen el significado de la violencia en esta época para un grupo de académicos y profesores universitarios, pero también plantea lo que está sucediendo en el escenario universitario. Esta situación, para muchos, puede ser percibida como normal o natural, porque forma parte de las relaciones cotidianas entre los individuos, y para otros es descubrir el mundo invisible, del cual se percatan hasta el momento en que reflexionan en ello.

Algunos profesores manifestaron que en su universidad no se presenta la violencia, pero, revisando las relaciones –sobre todo entre los y las estudiantes–, se aprecian las formas de competir por el poder, donde ambos sexos no se miran como humanos, sino como enemigos. La violencia se manifiesta, cada vez más, como una forma de contravenir al respeto y a la dignidad. Son más recurrentes las expresiones en charlas, coloquios, salas de debates, etc., sobre que la violencia se encuentra naturalizada; esto preocupa y detona nuestros quehaceres de investigación.

Para tomar medidas de acción, primero necesitamos saber qué entienden los académicos por violencia. Sabemos que se requiere de una cultura para la mejor convivencia, pero también se debe reconocer la dificultad de ver cambios inmediatos en una sociedad conmocionada como la nuestra. Primero, para comprender lo que se piensa de la violencia, debemos acercarnos al conocimiento de sentido común, y dar cuenta con precisión de las representaciones que actualmente se construyen sobre la violencia y nos orientan a comprender por qué los grupos sociales, como el universitario, actúan de una determinada manera.

Representación social

Para este estudio, nos apoyamos en la teoría de las representaciones sociales de Moscovici (1961), que nos dice que se desprenden del conocimiento de sentido común. Las define como un corpus organizado de lenguajes e imágenes inacabadas que producen determinadas acciones en los individuos que las poseen (Moscovici, 1961). Propone el abordaje del “proceso” de cómo son construidas y el análisis de su “contenido”; las explica como una forma de conocimiento que tiene el individuo sobre un objeto en un contexto social específico.

La representación social es el

estudio del pensamiento ingenuo del sentido común. La identificación de la visión de mundo que los individuos llevan en sí y utilizan para actuar o tomar posición es reconocida como indispensable para entender la dinámica de las interacciones sociales y aclarar los determinantes de las prácticas sociales (Abric, 1994, p. 12).

Abric (1994) asume una posición más estructural. Concibe que toda representación está organizada alrededor de un núcleo central, el cual determina la significación y la organización de la propia representación. Este núcleo central garantiza dos funciones esenciales: a) generadora, que transforma la significación de los otros elementos constitutivos de la representación; y b) organizadora, conocida como el núcleo central.

Lo que buscamos en este documento es mostrar los procesos que intervienen en la construcción de las representaciones sociales (RS) sobre la violencia, y las dimensiones desde las cuales permiten identificar la estructura y contenido a partir de la teoría del núcleo central y el sistema periférico (Abric, 1994) que un conjunto de profesores de la UV han elaborado a partir de su propia experiencia escolar (Dubet y Martucelli, 1998).

Es importante mencionar que, en los estudios de RS, se emplean los términos de “núcleo central” y “elementos periféricos”, ya que “toda representación social está organizada alrededor de un núcleo central” (Abric, 1994, p. 20). Específicamente, los elementos periféricos son los que se organizan alrededor del núcleo central; están en relación directa con él, con su valor y con su función; están determinados por el núcleo y dan un sentido a la estructura del núcleo.

El planteamiento de análisis del núcleo central y los elementos periféricos fue propuesto en la Universidad de Aix Provence, Francia, y por la Escuela de Psicología Social del país vasco. Los elementos constituyen lo esencial del contenido de la representación (Jodelet y Guerrero, 2000), su lado más accesible, pero también lo más vivo y concreto. Abarcan informaciones retenidas, seleccionadas e interpretadas; juicios formulados al respecto del objeto y su entorno; la representación se encuentra permeada de estereotipos o creencias. Estos elementos se encuentran jerarquizados y “desempeñan un papel importante en la concreción del significado de la representación, más distantes de él ilustran, aclaran, justifican esta significación” (Abric, 1994, p. 30).

El individuo ordena, organiza y jerarquiza el contenido de la representación. Le asigna un valor de importancia y atribuye un significado. El propósito es reflejar los contenidos centrales de la representación social, y de sus respectivos elementos periféricos (jerarquizados) que la integran. Estos últimos elementos “son un conjunto organizado de cogniciones relativas a un objeto, compartidas por los miembros de una población homogénea respecto a ese objeto” (Flament, 1994, p. 37).

La violencia

La violencia es una característica del ser humano. Se dice que, como acción o llevada a la práctica, no es frecuente. Sin embargo, la mayoría de los seres la ejercen en muchas formas y maneras (directa e indirecta), con gestos, palabras, actos, etc., con la intención de dañar y destruir. La violencia es un comportamiento adquirido y controversial (Blair, 2009). Es una respuesta a uno o una serie de estímulos. En su manifestación, admite pluralidad de las formas y matices (Baselga y Urquijo, 1974).

Autores como Arendt (2005), Harris (2006), Póo y Vizcarra (2008), Kaplan (2016), Dorantes (2017), y Bodelón (2010) han efectuado investigaciones sobre la violencia. Ellos precisan que la violencia es cuando los individuos han sido expuestos a situaciones dañinas, de agresividad, de maltrato, de humillación; de violencia física, psicológica, emocional, económica, verbal, de pareja, laboral, docente, de género e intrafamiliar. En los actos de violencia predominan tres actores: los victimarios, las víctimas y los testigos. La violencia manifiesta posiciones jerárquicas entre los individuos, donde la lucha del poder se manifiesta en el control y dominio del otro. La domesticación de la violencia han sido elementos determinantes para la conformación de la sociedad, tanto de la violencia que surge del entorno como de la que surge de la naturaleza humana... La propagación de la violencia entonces, puede interpretarse como la respuesta a la pérdida de solidaridad en las sociedades. (Kaplan, 2016, p. 216). La violencia es el resultado de la acción recíproca y compleja de los factores individuales, relacionales, sociales, culturales y ambientales, reflejándose en la pareja; pero, sobre todo:

es el uso deliberado de la fuerza física o el poder, ya sea en grado de amenaza o efectivo contra uno mismo, otra persona o un grupo o comunidad, que cause o tenga muchas probabilidades de causar lesiones, muerte, daños psicológicos, trastornos del desarrollo o privaciones (Organización Mundial de la Salud, 2002, p. 5).

Por su parte, Arendt (2005) precisa que la violencia tiene como fin causar un mal o un daño, sin siquiera identificar su gravedad en el individuo ni las consecuencias que pueda generar en la sociedad. Lo más grave es que “en los últimos años los centros educativos, comunidades y padres de familia de todo el mundo han empezado a reconocer las conductas de acoso y agresión entre escolares” (Harris y Petrie, 2006, pp. 11-12); asimismo agrega “la violencia escolar puede afectar negativamente la experiencia escolar del agresor, del agredido y del espectador” (Harris y Petrie, 2006, p. 15), por lo que se debe de indagar sus causas.

A nivel internacional la política internacional denominada Educación 2030. Incheon. Declaración de Incheon y Marco de Acción. Hacia una educación inclusiva y equitativa de calidad y un aprendizaje a lo

largo de la vida para todos, establece “apoyar políticas planes y contextos de aprendizaje en que se tengan en cuenta las cuestiones de género así como incorporar estas cuestiones en la formación de docentes, los planes de estudios, y a eliminar la discriminación y la violencia por motivos de género en las escuelas” (Declaración de Incheon y Marco de Acción ODS 4 Educación 2030, P, 8); incluye a todos los niveles y modalidades educativas y por supuesto a las universidades de nuestro país México y a 160 países del mundo que se comprometieron con la meta como una nueva visión de la educación para los próximos 15 años.

METODOLOGÍA DE TRABAJO

Para abordar este estudio, iniciamos con el análisis de una pregunta abierta planteada a un grupo de profesores universitarios, que asistieron a un curso de RS. La pregunta uno, planteada en el estudio, se efectuó a un total de 21 académicos de la UV sobre la violencia que asistieron al taller “Representaciones sociales”, brindado por Ana María Silvana de Rosa, de la Universidad Sapienza de Roma y European PhD on Social Representations & Communications, en el marco del V Coloquio Nacional de Representaciones Sociales en la Universidad Veracruzana, a cargo del Centro de Estudios de Opinión y Análisis (CEOA) y el Programa de la Especialización en Estudios de Opinión (EEO), durante octubre de 2015 en las instalaciones de la Unidad de Servicios Bibliotecarios (USBI), región Xalapa. Este taller se brindó en italiano, y requirió de la traducción simultánea al español, a cargo de la maestra en Antropología Elisa Manno, de la UV, quien cuidó de manera profesional que las expresiones, modismos y términos teóricos fueran impecables para el público de académicos asistentes.

Los académicos participantes provinieron de las diversas entidades que conforma a la UV: CEOA, Centro de Enseñanza de la Lengua Extranjera (CELE), Dirección General del Área de Humanidades, Instituto de Psicología y Educación, Instituto de Investigaciones en Educación (IIE), Maestría en Desarrollo Humano, Maestría en Antropología, Maestría en Investigación Educativa, Especialización en Estudios de Opinión, Facultad de Ciencias Administrativas y Sociales, Facultad de Sociología, Facultad de Idiomas (Lengua Francesa e Inglesa) y la Facultad de Estadística; así como dos académicos de la Universidad Autónoma Metropolitana (UAM) y la Benemérita Universidad Autónoma de Puebla (BUAP); y uno de la Universidad de Brest-Francia. Todos ellos, interesados en aprender de los métodos y técnicas de investigación para el estudio de las representaciones sociales, pero también, sensibles e interesados ante el tema de la violencia.

Para comprender y conocer las RS sobre la violencia de los académicos de la UV, identificamos sus elementos importantes: el núcleo central y los elementos periféricos. La técnica de investigación empleada fue la asociación libre Abrie (1994) “permite reducir la dificultad o los límites de la expresión discursiva. Consiste, a partir de un término inductor (o de una serie de términos), en pedir al sujeto que produzca todos los términos, expresiones o adjetivos que se le presentan al espíritu. El carácter espontáneo-permiten el acceso, mucho más rápido y fácil que una entrevista, a los elementos que constituyen el universo semántico del objeto estudiado. La asociación libre permite actualizar elementos implícitos o latentes que serían ahogados o enmascarados en las producciones discursivas..., permiten el acceso a los núcleos figurativos de la representación” (p. 59). Para ello, planteamos una premisa de investigación: Pregunta 1: Escriba 6 palabras o expresiones que le vienen a la mente inmediatamente después de pronunciar la palabra violencia.

Una vez obtenida la información, se capturó en Word, revisó y corrigió la ortografía de las respuestas obtenidas para el posterior análisis. Los datos se guardaron en un bloc de notas con codificación UTF-8, y desde el *software* Iramuteq (Interfaz de R para el Análisis Multidimensional de los Textos y Cuestionarios) se realizó un gráfico de nubes y un análisis de similitud. En este análisis, pudimos observar las representaciones sociales que los profesores tienen ante la pregunta.

Iramuteq es un *software* libre que permite realizar análisis de textos, como preguntas de respuesta abierta de cuestionarios, entrevistas, etc. Al realizar el análisis, elegimos la opción de lematización en Iramuteq, este “proceso reduce las palabras principales a sus raíces, deviniendo así ‘morfemas lexicales’” (Molina-Neira, 2017, p. 14) para facilitar la interpretación de los resultados. Dicho proceso toma en cuenta la raíz de las palabras, sin tomar en cuenta el género o la cantidad.

La nube de palabras agrupa y organiza las palabras gráficamente en función de su frecuencia. Permite visualizar de manera nítida los datos a partir de las respuestas de los informantes.

Efectuamos un análisis de similitud (ADS), que es el enfoque basado en las propiedades de conexión del corpus completo, sin tener en cuenta las unidades de contexto específicas ni los sujetos. Identificamos las coocurrencias entre las palabras según sus conexiones en el texto, ayudando a identificar la estructura del contenido del corpus textual gracias a su visualización en forma de gráfico, que ilustra el contenido de la representación social del objeto estudiado y su organización interna, sus partes comunes y especificidades (Marchand y Ratinaud, como se citó en Molina-Neira, 2017, p. 5).

En el proceso de análisis logramos identificar los núcleos representacionales de los sujetos (profesores de la UV), ya que el programa registra los núcleos semánticos detectados por coocurrencias de palabras,

logrando trabajar bajo el paraguas del enfoque estructural de las representaciones sociales; incluso, obteniendo los datos a partir de entrevistas y no de técnicas propias de la investigación experimental (Abric, 1994; Doise y Lorenzi, 2005; y Catorina y Barreiro, 2005).

En el seminario de violencia participaron un total de 21 académicos de la UV, invitados a un curso sobre RS, a quienes se les pidió escribir seis palabras o expresiones que les vinieran a la mente inmediatamente después de pronunciar la palabra violencia. Buscábamos saber cómo es que representan la violencia y, con ello, dar cuenta de qué es lo que entienden cuando se habla del término. Los académicos anotaron las palabras en una ficha de papel.

Las respuestas se clasificaron en orden de importancia, se capturaron y se analizaron por medio del *software* Iramuteq. La intención fue conocer la representación social, pero también el núcleo central y los elementos periféricos que la componen.

RESULTADOS

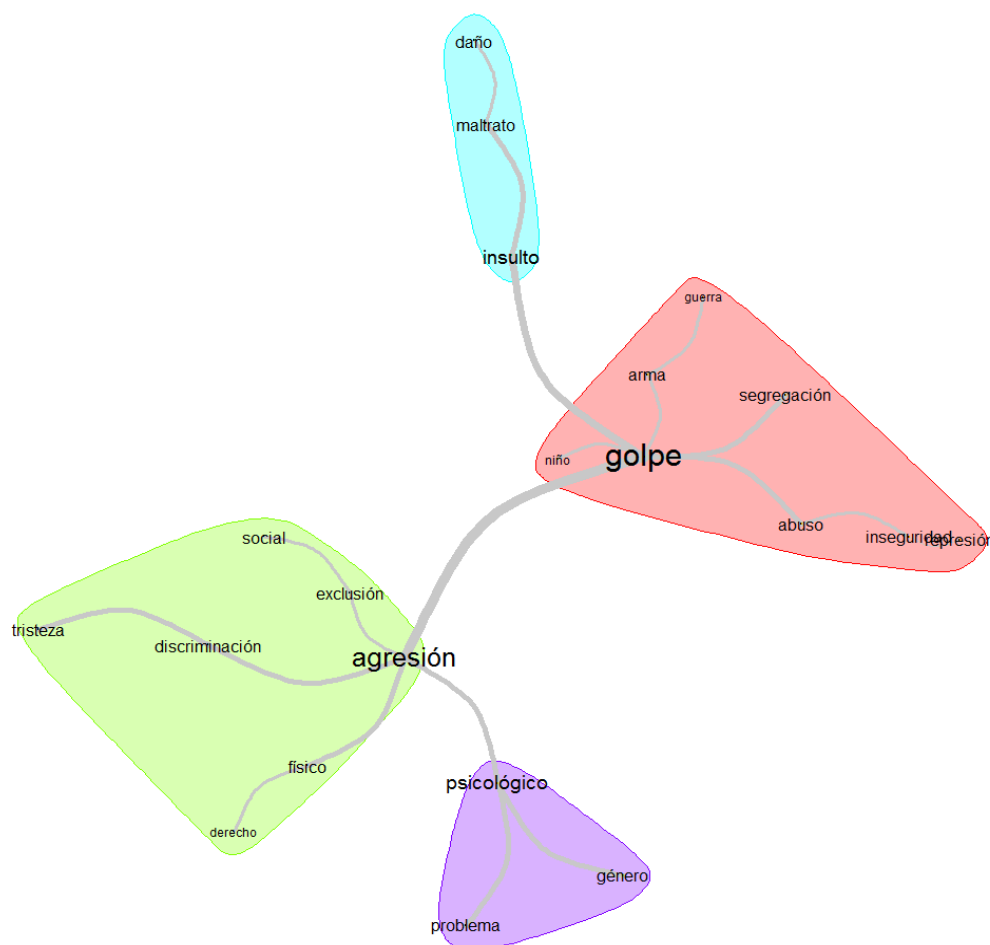
Observamos que 21 académicos de la UV representan en sus mentes a la violencia con diferentes palabras; pero “golpe” conforma al núcleo central de la representación social, pues es el dato más duro y frecuente que destaca cuando se habla de violencia. Veamos el gráfico de nube de palabras:

Gráfico 1. Nube de palabras



Sin embargo, observamos que las palabras agresión, maltrato, segregación, niños, social, inseguridad, abuso, tristeza, psicológico, físico, insulto, problema, derecho, exclusión, problema, daño, guerra, género y discriminación conforman los elementos periféricos. También se encuentran presentes en el gráfico de nube de palabras.

Gráfico 2. Gráfico de similitud



Observamos que el núcleo central de la representación social elaborada por los académicos de la UV sobre violencia es “golpe”, y los elementos periféricos son “agresión”, “insulto” y “psicológico”.

Los profesores universitarios opinaron que la violencia está relacionada con los golpes. Los golpes son entre niños, pero también implican la guerra y en esta se utilizan armas que generan violencia. Entre los actos que se presentan dentro de este fenómeno son daño, insulto y maltrato. El abuso de autoridad genera inseguridad y represión ante este fenómeno y segregación por parte de las personas que viven la violencia. Los golpes generan problemas psicológicos y un golpe es una agresión que daña al físico. Reconocen que en los “derechos” hay herramientas para defenderse; conciben que la violencia es resultado de la “discriminación, exclusión y de los problemas sociales”. El escenario de la violencia les genera tristeza. El género también es un tema que está en controversia.

Póo y Vizcarra (2008) señalan que el impacto de la violencia psicológica es de largo plazo, pues incluye actos de coerción, manipulación y presión, concentrando los eventos de este tipo de violencia en actitudes de control y subestimación. Quien ejerce esta forma de maltrato tiene la intención de humillar y degradar

al otro, buscando que se sienta mal al decirle, por ejemplo, palabras hirientes, usando la mordacidad, bromas hirientes, la mentira, la ridiculización, el sarcasmo, el silencio, el aislamiento e, incluso, las amenazas de ejercer otras formas de violencia.

CONCLUSIONES

Gracias a los *softwares* estadísticos, logramos conocer las representaciones de los académicos de la Universidad Veracruzana. Identificamos el núcleo central de la representación social compartida; obtuvimos la palabra “golpe”, que es la más frecuente, es decir, cuando se les habla de la representación social de la violencia, lo primero que se les viene a la mente es la violencia física, siendo la que más perciben u observan en el entorno universitario y en el contexto de las aulas o salones de la UV. También reconocieron que la violencia es “agresión”, que se articula a insultos, los cuales surgen ante la agresión y que estos, a su vez, causan un daño psicológico a los estudiantes universitarios.

A pesar de que no mencionan más tipos de violencia, logramos identificar, a partir del Gráfico de similitud de palabras, “violencia es la agresión, el insulto y el daño psicológico”. También destacaron la palabra agresiones y la discriminación del físico de alguna persona. Las RS de los académicos corresponden a lo que ellos, desde su mirada muy particular, perciben en el contexto universitario.

Como profesores, como académicos, deben apoyar a mejorar los ambientes entre los estudiantes, deben promover el cambio desde las aulas, son ellos quienes pueden promover nuevas formas de relacionarse entre estudiantes y profesores. Hablamos de nuevas prácticas más respetuosas y tolerantes, ya que la Universidad Veracruzana es el espacio más importante, social y profesionalmente, donde debe imperar la cultura de respeto. Eso es lo que nos debe distinguir como “comunidad universitaria”.

Se deben desmontar los pilares de la dominación, del poder y de la discriminación. La violencia en cualquiera de sus manifestaciones: física, psicológica o emocional, de género, sexual, *cyberbullying*, debe ser erradicada de las aulas universitarias. Conocer las RS de los académicos universitarios, el núcleo central y sus elementos periféricos nos invita a dar cuenta de que la violencia aún impera en las prácticas universitarias. Contribuyamos cada uno de nosotros a cambiar dichas prácticas, y promulguemos, sembremos y seamos ejemplo de prácticas de respeto, la tolerancia y la valoración de la diversidad, ya que forman parte de nuestros derechos humanos.

Finalmente, podemos decir que el estudio de las RS nos permite reconocer que la violencia genera afectaciones psicológicas en el individuo. Por otra parte, la estadística en las investigaciones es una excelente herramienta de investigación, que nos permite tener mayor precisión, exactitud y rigurosidad en

el manejo de los datos. La estadística nos invita a comprender las realidades de manera más precisa. Hoy logramos captar con mayor exactitud cuáles son las formas de pensar de un grupo social de académicos universitarios ante “la violencia”, como un fenómeno real que aún se percibe en los entornos universitarios, que forma parte de nuestras vidas cotidianas y que caracteriza a la sociedad en la cual estamos inmersos.

REFERENCIAS

- Abramovay, M. (Coord.) (2006). *Cotidiano das escolar: entre violências*. Brasília, Brasil: UNESCO.
- Abric, J. C. (1994). *Prácticas sociales y representaciones*. México: Ed. Coyoacán.
- Almeida, C., y Gómez, M. (2005). *Las huellas de la violencia invisible*. Barcelona, España: Ariel.
- Arendt, H. (2005). *Sobre la violencia*. Madrid, España: Taurus.
- Arisó, O., y Mérida, R. M. (2010). *Los géneros de la violencia. Una reflexión queer sobre la violencia de género*. Barcelona-Madrid, España: Egales.
- Baselga, E., y Urquijo, S. (1974). *Sociología y violencia. Actitudes, universitarios*. Bilbao, España: Universidad de Deusto.
- Blair, E. (2009). *Aproximación teórica al concepto de violencia: avatares de una definición*. Medellín, Colombia: Universidad de Medellín.
- Blaya, C. (2006). *Violences et maltraitances en milieu scolaire*. Quebec, Canadá: Armand Colin.
- Bodelón, E. (2010). *Violencia de género y las propuestas de los sistemas penales*. Buenos Aires, Argentina: Didot.
- Castro, A., y Reta, C. (2013). *Bullying blando, bullying duro y cyberbullying. Nuevas violencias y consumos culturales*. Argentina: Homo Sapiens.
- Catorina, J. A., y Barreiro, A. (2005). Dos versiones del sentido común: las teorías implícitas y las representaciones sociales. En J. A. Castorina (Coord.), *Construcción conceptual y representaciones sociales. El conocimiento de la sociedad* (pp. 215-238). Argentina: Minó y Dávila.
- De Rosa, A. S. (2005). A “Rede Associativa”: uma técnica para captar a estrutura, os conteúdos, e os índices de polaridade, neutralidade e estereotipia dos campos semânticos relacionados com a Representações Sociais. In A. S. Paredes Moreira (Ed.), *Perspectivas Teorico-metodológicas em Representações Sociais* (pp. 61-127). Brasil: Universitária-UFPB, João Pessoa.

- Doise, W., y Lorenzi, F. (2005). *Representaciones sociales y análisis de datos*. México: Instituto de Investigaciones Dr. José María Luís Mora.
- Dorantes, J. J. (2017). Las representaciones sociales de los estudiantes sobre la violencia. Caso Universidad Veracruzana. En M. E. Ruíz Libreros, *Complejidad, innovación y sustentabilidad. Experiencias educativas* (pp. 163-179). Xalapa, Veracruz, México: CÓDICE/Taller editorial.
- Dubet, F., y Martucelli, D. (1998). *En la escuela. Sociología de la experiencia escolar*. Buenos Aires, Argentina: Losada.
- Educación 2030. *Incheon. Declaración de Incheon y Marco de Acción. Hacia una educación inclusiva y equitativa de calidad y un aprendizaje a lo largo de la vida para todos*. (21 de mayo 2015). República de Corea: UNESCO, UNICEF, Banco Mundial, UNFPA, PNUD, ONU Mujeres y el ACNUR.
- Flament, C. (1994). Estructura, dinámica y transformación de las representaciones sociales. En J. C. Abrie (Coord.), *Prácticas sociales y representaciones*. México: FCE.
- Furlán, A. (2005). Problemas de indisciplina y violencia en la escuela. *Revista Mexicana de Investigación Educativa*, 10(26), 631-639.
- Furlán, A., Pasillas, M. A., Spitzer, T. C., y Gómez, A. (Coords.). (2010). *Violencia en los centros educativos. Conceptos, diagnósticos e intervenciones*. Buenos Aires, Argentina: Noveduc.
- Grettiez, X. (2008). *Las formas de la violencia*. Buenos Aires, Argentina: Waldhunter.
- Guillotte, A. (2003). *Violencia y educación. Incidentes, incivildades y autoridad en el contexto escolar*. España: Amorrortu.
- Harris, S., y Petrie, G. (2006). *El acoso en la escuela. Los agresores, las víctimas y los espectadores*. Barcelona, España: Paidós.
- Jacorzynski, W. (2003). *Estudios sobre la violencia. Teoría y práctica*. México: Centro de Investigaciones y Estudios Superiores en Antropología Social (CIESAS).
- Jodelet, D., y Guerrero, T. (Eds.). (2000). *Develando la cultura: estudios en representaciones sociales*. México: UAM.
- Johnson, W. D., y Johnson, T. R. (2010). *Cómo reducir la violencia en las escuelas*. Argentina: Paidós Educador.
- Kaplan, C. V. (2016). "El lenguaje es una piel. Género, violencia y procesos civilizatorios." En Kaplan, C. V. (Editora). *Género es más que una palabra. Educar sin etiquetas*. Buenos Aires: Miño y Dávila. pp. 211-223.

- Molina-Neira, J. (2017). *Tutorial para el análisis de textos con el software Iramuteq*. Barcelona, España: Grupo de Investigación DHIGES, University of Barcelona.
- Morales, R. T., Serrano, B. C., Miranda, G. D., y Santos L. A. (2014). *Cyberbullying, acoso cibernético y delitos invisibles. Experiencias psicopedagógicas*. México: Universidad Autónoma del Estado de México.
- Moscovici, S. (1961). *El psicoanálisis, su imagen y su público*. Buenos Aires, Argentina: Huemul.
- Novo, M., y Seijo, D. (2009). Aproximación psicosocial a la violencia de género: Aspectos introductorios. En F. Fariña, R. Arce y G. Buela-Casal, *Violencia de género. Tratado psicológico y legal* (pp. 63-73). España: Biblioteca Nueva.
- Sanmartín, J. (2010). *La violencia y sus claves*. Barcelona, España: Ariel.
- Sofsky, W. (2006). *Tratado sobre la violencia*. Madrid, España: Abad.
- Torres, M. F. (2005). *La violencia en casa*. México: Paidós.

Referencias electrónicas

- Organización Mundial de la Salud. (2002). *Informe mundial sobre la violencia y la salud*. Recuperado de http://apps.who.int/iris/bitstream/handle/10665/67411/a77102_spa.pdf;jsessionid=CCBCA176B51CB39677616C478E2B090D?sequence=1.
- Póo, A. M., y Vizcarra, B. (2008). Violencia de pareja en jóvenes universitarios. *Terapia Psicológica*, 26(1), 81-88. Recuperado de https://scielo.conicyt.cl/scielo.php?script=sci_arttext&pid=S0718-48082008000100007
- Rojas-Solís, J. L. (2013). Violencia en el noviazgo y sociedad mexicana posmoderna. Algunos apuntes sobre la figura del agresor y las agresiones bidireccionales. *Uaricha*, 10(22), 1-19. Recuperado de http://www.revistauaricha.umich.mx/Articulos/uaricha_1022_001-019.pdf
- Universidad Veracruzana. (2013). *Programa de Trabajo Estratégico 2013-2017 Tradición e Innovación*. Recuperado de <https://www.uv.mx/documentos/files/2019/05/Programa-de-Trabajo-Estrategico-version-para-pantalla.pdf>

EFFECTIVIDAD DEL MECANISMO DE TRANSMISIÓN DE LA POLÍTICA MONETARIA EN MÉXICO (2020-2023)

Christopher Cernichiaro Reyna⁵, Edson Valdés Iglesias⁶ Roberto Gallardo Del Ángel⁷

Universidad Autónoma Metropolitana, Universidad Veracruzana
ccernichiaro@correo.xoc.uam.mx, edvaldes@uv.mx, rogallardo@uv.mx.

RESUMEN

El objetivo de esta investigación es verificar el cumplimiento del canal de transmisión de la política monetaria en México. Para ello, se estima un modelo de vectores autorregresivos estructurales con datos mensuales de febrero de 2020 a junio de 2023. Los resultados muestran que, durante la pandemia, la tasa de interés nominal responde positivamente a incrementos en la inflación y a la depreciación del tipo de cambio real. Además, un aumento en la tasa de interés disminuye la demanda agregada, probablemente debido a los fuertes incrementos en la tasa durante la pandemia. Sin embargo, la inflación no disminuye como respuesta a estos incrementos, posiblemente debido a los choques exógenos inflacionarios ocurridos durante la pandemia.

Palabras clave: Canal de transmisión de la política monetaria; Tasa de interés nominal.

INTRODUCCIÓN

Los bancos centrales a nivel mundial enfrentan desafíos complejos debido a choques externos e internos, exacerbados por las condiciones económicas globales que afectan variables macroeconómicas clave como las tasas de política monetaria, los tipos de cambio, la demanda, la oferta y la inflación. Países como Italia y España han logrado estabilizar sus economías y controlar la inflación, mientras que naciones como Canadá y Chile enfrentan presiones inflacionarias persistentes al priorizar la estabilización del producto. Polonia, tras una rápida recuperación económica post-COVID-19, experimentó un declive en el crecimiento, acompañado de alta inflación y baja actividad económica.

⁵ Universidad Autónoma Metropolitana, Unidad Xochimilco [ccernichiaro@correo.xoc.uam.mx].

⁶ Facultad de Economía, Universidad Veracruzana. [edvaldes@uv.mx]

⁷ Facultad de Economía, Universidad Veracruzana. [rogallardo@uv.mx]

En Italia, el canal de transmisión de la política monetaria ha sido consistentemente efectivo antes y después de la pandemia. Según Dallari y Riba (2023), entre 1999 y 2016, la tasa de política controlaba la inflación, aunque a costa de mayores tasas de desempleo. En 2023, la política monetaria se incrementó para mantener ancladas las expectativas de inflación, a pesar de los efectos adversos sobre la actividad económica, particularmente en la inversión (IMF, 2023a).

En España, el funcionamiento del canal de transmisión se alinea con las expectativas teóricas. Dallari y Riba (2023) no encontraron un impacto estadístico significativo de la política monetaria sobre la inflación durante la pandemia, aunque el mecanismo fue efectivo. El FMI (2023c) reporta que el producto está casi en su nivel objetivo, permitiendo que la tasa de interés nominal se concentre en estabilizar las expectativas de inflación.

En Canadá, la efectividad del canal de transmisión de la política monetaria ha variado antes y después de la pandemia, debido a múltiples choques exógenos que afectaron la demanda y la oferta. Ambler y Kronick (2021) señalaron que, durante la década de 2010, el enfoque de la política fue alinear la inflación con su objetivo, lo que llevó a aumentos en las tasas de interés, incluso en recesiones económicas. Hacia el final de la década, en 2018-2019, la política se centró en la estabilidad del PIB en medio de recesiones globales, a pesar de las presiones inflacionarias en curso. Azad et al. (2021) observaron que, durante la pandemia, las reducciones en las tasas de interés disminuyeron la inversión sin controlar eficazmente la inflación. Los desafíos en el control de precios pueden derivarse de la recuperación económica asimétrica y disrupciones en la oferta debido a conflictos globales (Rodríguez et al., 2023).

En Chile, la transmisión de la política monetaria antes y después de la pandemia ha sido consistente, aunque los choques de demanda y oferta internos han exacerbado la brecha inflacionaria desde el inicio de COVID-19. Antes de la pandemia, el banco central estabilizó eficazmente la inflación y promovió el crecimiento bajo un régimen de control estricto. Sin embargo, perturbaciones sociales en 2019 y una caída del 15% en el producto en abril de 2020 desestabilizaron tanto la inflación como el crecimiento económico. La política monetaria se centró en estabilizar el producto hasta la recuperación económica en 2021. No obstante, el crecimiento económico típicamente trae consigo una mayor inflación, intensificada por choques externos e internos como el aumento de la liquidez y los costos energéticos (Costa, 2023).

En Polonia, la efectividad de la política monetaria ha disminuido tras la COVID-19. Dallari y Riba (2023) observaron que, antes de la pandemia, la política monetaria contractiva reducía efectivamente la inflación y el crecimiento económico. Polonia enfrenta alta inflación y un crecimiento económico débil, a pesar de

las tasas de política aumentadas. De 2020 a octubre de 2021, la tasa de interés nominal escaló de cero a 6.75%, disminuyendo el consumo y la inversión sin controlar eficazmente la inflación (IMF, 2023b).

Ante estas incertidumbres globales, este estudio evalúa la dinámica del canal de transmisión de la política monetaria en la economía mexicana post-pandemia mediante Modelos Autorregresivos Vectoriales Estructurales (SVAR), analizando cómo los aumentos en las tasas de interés nominales responden a incrementos en la inflación, el producto y el tipo de cambio real, y cómo estos factores interactúan en un contexto de inestabilidad económica.

La sección siguiente delimita el canal de transmisión de la política monetaria canónica (Galí y Monacelli, 2005). A continuación, se describe la metodología SVAR empleada para evaluar el canal de transmisión de la política monetaria. Después, se discuten los resultados, interpretándolos según los criterios de los autores. Finalmente, se exponen las conclusiones.

MODELO TEÓRICO

Varias teorías proponen cómo funciona el canal de transmisión de la política monetaria. Un modelo destacado es el Modelo Nuevo Keynesiano, derivado de la teoría neoclásica, que sugiere fallas de mercado que justifican la intervención gubernamental. Clarida et al. (1998) lo propusieron inicialmente para una economía cerrada, pero Galí y Monacelli (2005) lo extendieron a una economía abierta con tipo de cambio flotante. Este modelo, con fundamentos microeconómicos, ha sido influyente durante casi dos décadas y sigue proporcionando lecciones, especialmente durante la pandemia (IMF, 2023 a,b,c).

El marco teórico empleado en este artículo es el modelo de economía abierta Nuevo Keynesiano de Galí y Monacelli (2005). Se incorporan los contagios de COVID-19 en el modelo debido a su relevancia empírica en la explicación del canal de transmisión de la política monetaria durante la pandemia (Eichenbaum et al., 2021; Guerrieri, 2022). Esta sección proporciona una visión general del modelo teórico. La estructura macroeconómica de Galí y Monacelli (2005) se caracteriza por la función IS, la curva de Phillips y la Regla de Taylor, en las ecuaciones 1, 2 y 3, respectivamente.

$$\check{y}_t = E_t \check{y}_{t+1} - \frac{1}{\sigma_\alpha} (i_t - E_t \pi_{H,t+1} - r_t^n) + \varepsilon_{1,t} \quad (1)$$

$$\pi_{H,t} = \beta E_t \pi_{H,t+1} + k_\alpha \check{y}_t + \varepsilon_{2,t} \quad (2)$$

$$i_t = \rho + \phi_\pi \pi_t + \phi_y \check{y}_t + \varepsilon_{3,t} \quad (3)$$

Donde $\check{y}_t$ es la brecha del producto, $E_t\check{y}_{t+1}$ es la brecha del producto esperada en el periodo $t+1$, i_t es la tasa de interés nominal, r_t^n es la tasa de interés natural, $\varepsilon_{1,t}$ es un choque de demanda agregada, $\pi_{H,t}$ es la inflación doméstica, $\pi_{H,t+1}$ es la inflación esperada para $t+1$, $\varepsilon_{2,t}$ es un choque de oferta agregada, y $\varepsilon_{3,t}$ es un choque de política monetaria. Mientras que $\sigma_\alpha, \beta, k_\alpha, \rho, \phi_\pi, \phi_y$ son parámetros positivos.

El canal de transmisión de la política monetaria teórico es el siguiente: Ante un aumento exógeno de la tasa de política monetaria, la demanda agregada disminuye porque los hogares posponen el consumo, la inversión se encarece y el tipo de cambio real se aprecia, reduciendo las exportaciones netas. Los productores responden disminuyendo la producción, lo que reduce los costos marginales y los precios establecidos, por lo tanto, la inflación disminuye.

El modelo descrito sigue siendo relevante para explicar el canal de transmisión de la política monetaria, como lo demuestran los informes del FMI (2023 a,b,c) que utilizan una versión similar. Sin embargo, se propone incorporar los contagios de COVID-19 debido a su reacción con diversos choques, como el cierre de sectores intensivos en contacto, restricciones comerciales y interrupciones en la oferta laboral etcétera (Eichenbaum et al., 2021; Guerrieri et al., 2022).

MATERIALES Y MÉTODOS

Esta sección analiza los datos empleados en las estimaciones, alineando las variables elegidas con el modelo teórico presentado. Las variables utilizadas se basan en metodologías de medición establecidas en estudios previos, como los de Cermeño et al. (2012), León y Martínez (2013), Loría y Ramírez (2011) y Ros (2015), entre otros. Los datos son específicos para México y cubren el periodo de febrero de 2020 a junio de 2023. La recopilación de datos se realizó a partir de fuentes oficiales, como se detalla en la Tabla 1. Las variables de PIB y tipo de cambio real se presentan en precios constantes de 2013, mientras que la tasa de interés se expresa en términos nominales.

Tabla 1. Variables y fuentes de información.

Tasa de interés nominal	Banco de México (2023a)
IGAE	INEGI (2023a)
Inflación	INEGI (2023b)
Tipo de Cambio Real	Banco de México (2023b)
Contagios de COVID-19	CONAHCYT (2023)

Fuente: Elaboración propia.

TRANSICIÓN DE ESTADÍSTICA A CIENCIA DE DATOS

La tasa de política monetaria se aproxima utilizando el rendimiento de los bonos gubernamentales a un mes, específicamente los CETES a 28 días. Para evaluar la evolución del PIB, se emplea el Indicador Global de Actividad Económica (IGAE), el cual proporciona una visión del desempeño a corto plazo del sector real de la economía.

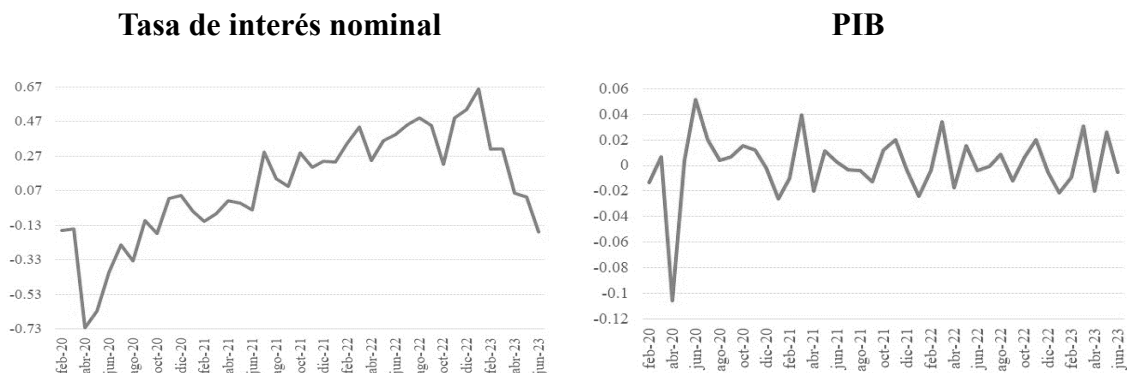
La tasa de inflación se mide utilizando el componente subyacente del Índice Nacional de Precios al Consumidor, que captura los cambios típicos en los precios de una canasta predeterminada de bienes y servicios adquiridos por los hogares urbanos, excluyendo deliberadamente los precios sujetos a volatilidad o control gubernamental. Además, se considera el tipo de cambio real, calculado como la tasa efectiva real promedio a lo largo de 49 países. Finalmente, el total de contagios de COVID-19 representa el agregado de infecciones confirmadas en México. La Tabla 2 muestra las estadísticas descriptivas de las variables presentadas en la Figura 1. Cada variable se expresa como la primera diferencia de su logaritmo, excepto la tasa de interés nominal, que se presenta como la primera diferencia en niveles.

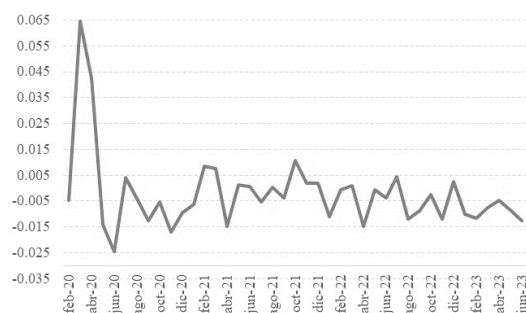
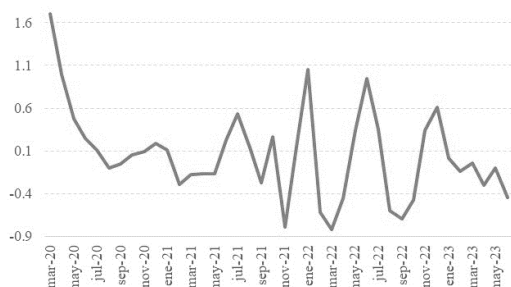
Tabla 2. Estadísticas descriptivas.

	Tasa de interés nominal	PIB	Inflación	Tipo de cambio real	Contagios de COVID-19
Observaciones	41	41	41	41	41
Media	0.098439	0.000607	0.491707	-0.002205	0.056406
Mediana	0.095000	-0.00064	0.480000	-0.004630	0.014383
Máximo	0.656500	0.051622	0.800000	0.064660	1.702041
Mínimo	-0.72050	-0.10579	-0.08000	-0.024591	-0.816232
Desv. Est.	0.310544	0.024606	0.186224	0.015032	0.525966

Fuente: Elaboración propia.

Figura 1. Tasa de interés nominal, PIB, inflación, tipo de cambio real y contagios de COVID-19.



Inflación**Tipo de cambio real****Contagios de COVID-19**

Fuente: Elaboración propia.

La tasa de interés nominal disminuyó desde la segunda mitad de 2019 hasta junio de 2021, reduciéndose de 8.25% a 4%, probablemente para mitigar la contracción del PIB en 2019, que se debilitó un 0.3% debido a contracciones en el sector manufacturero en las regiones central y norte del país (Banco de México, 2020). En abril de 2020, la actividad económica industrial cayó un 21.6% debido al aumento recurrente de casos de COVID-19, que pasaron de 100 en marzo a 988 en abril.

A pesar de que la actividad económica no se recuperó a niveles pre-pandemia, la tasa de interés nominal se focalizó en estabilizar la tasa de inflación a mediados de 2021. La tasa comenzó a incrementarse en julio de 2021 (4.3%), alcanzando el 11.15% en junio de 2023, con el objetivo de controlar la inflación, que aumentó casi cinco puntos porcentuales de enero de 2021 a diciembre de 2022, pasando de 3.83% a 8.34%. Para junio de 2023, la tasa de inflación se situó en 6.89%, aproximadamente tres puntos porcentuales más alta que en enero de 2021.

Cuando ocurrió una depreciación del tipo de cambio real en abril de 2020, la tasa de interés nominal disminuyó. Según Ros (2015), cuando el tipo de cambio real se deprecia, el Banco Central eleva la tasa

de interés para proteger al tipo de cambio de una mayor depreciación. Además, Rodríguez et al. (2023) señala que, al inicio de la pandemia, a pesar de los choques inflacionarios compensatorios, fue posible estabilizar la caída del producto, aunque se requería precaución debido a que la depreciación del tipo de cambio indicaba posibles presiones inflacionarias futuras.

En resumen, en esta sección se utilizaron indicadores económicos para analizar las variables relevantes de la política monetaria en México. El análisis estadístico descriptivo sugiere que la tasa de interés nominal alterna entre enfatizar la estabilidad de la inflación y del PIB, a pesar de que el único mandato legal del Banco de México es procurar el poder adquisitivo del peso. En los casos examinados, ya sea la tasa de inflación o el PIB se estabilizan, aunque no regresan a niveles previos a la intervención. Para profundizar en el desempeño empírico de las relaciones teóricas de la política monetaria, las secciones subsiguientes realizan análisis econométricos.

JUSTIFICACIÓN DE LA METODOLOGÍA

El uso de Modelos SVAR es particularmente adecuado para la investigación de la transmisión de la política monetaria en México por varias razones. En primer lugar, el SVAR permite identificar relaciones dinámicas entre múltiples variables económicas sin imponer restricciones superficiales (Das, 2019). Esto es especialmente relevante en el análisis de políticas monetarias, donde los efectos de las acciones del banco central pueden depender de condiciones macroeconómicas específicas, como inflación alta o producción baja.

En segundo lugar, Ouliaris et al. (2016) destacan que los modelos SVAR son eficaces para analizar las dinámicas de corto plazo, permitiendo capturar las interacciones entre variables como la tasa de interés, el producto, la inflación y el tipo de cambio para evaluar el cumplimiento del canal de transmisión de la política monetaria.

Por último, Stock y Watson (2020) señalan que los SVAR son herramientas robustas para evaluar los efectos de políticas económicas bajo incertidumbre, lo cual es fundamental dada la volatilidad introducida por la crisis sanitaria y las respuestas políticas subsecuentes. El uso del SVAR en este estudio permite no solo identificar cómo la política monetaria ha afectado variables clave, sino también cómo estas relaciones pueden haber cambiado debido a choques exógenos significativos.

RESULTADOS

Esta sección presenta evidencia que respalda la idoneidad de los datos para estimar un modelo SVAR sin relaciones espurias en el periodo de febrero de 2020 a junio de 2023. Inicialmente, se realizaron pruebas de estacionariedad para las variables individuales. Posteriormente, se discutió la determinación de la longitud de rezago apropiada para cada modelo VAR, seguida de pruebas de estabilidad y correlación serial. Luego de estos exámenes, se presentaron las Funciones de Impulso Respuesta.

Las variables se expresan como la primera diferencia del logaritmo, excepto la tasa de interés nominal, que se muestra como la primera diferencia en niveles. Esta diferenciación se realizó usando el valor del mismo mes del año anterior para comparar meses afectados por la pandemia, evitando contrastar un mes del año pandémico con uno del año no pandémico.

En cuanto a las pruebas individuales de estacionariedad, las pruebas Aumentada de Dickey Fuller (ADF), Phillips-Perron (PP) y Kwiatkowski-Phillips-Schmidt-Shin (KPSS) no muestran consenso sobre la estacionariedad de las variables en ambos periodos analizados (Tablas 3 y 4). Todas las pruebas indican algún nivel de estacionariedad para las variables, excepto para la tasa de interés nominal, donde solo la prueba KPSS sugiere una ligera indicación de estacionariedad.

Tabla 3. Pruebas de estacionariedad de ADF y PP.

Variable*	ADF		PP	
	Intercepto	Tendencia e intercepto	Intercepto	Tendencia e intercepto
Tasa de interés nominal	0.4145	0.5590	0.4656	0.5870
PIB	0.0000	0.0000	0.0000	0.0000
Inflación	0.4017	0.8101	0.0295	0.0323
Tipo de cambio real	0.0001	0.0001	0.0010	0.0021
Contagios	0.0001	0.0002	0.0000	0.0000

* Las celdas sombreadas indican el rechazo de la hipótesis nula de raíz unitaria, confirmando la estacionariedad de las series temporales. Fuente: elaboración propia.

Tabla 4. Prueba de estacionariedad KPSS.

	Tasa de int. nominal		PIB		Inflación		Tipo de cambio real		Contagios	
	Inter c.	Tendencia e interc.	Inter c.	Tendencia e interc.	Interc.	Tendencia e interc.	Interc.	Tendencia e interc.	Interc.	Tendencia e interc.
Estadístico de KPSS	0.57653	0.170536	0.273924	0.195375	0.43256	0.160483	0.306622	0.079766	0.456988	0.130646

Valores críticos :	1% nivel	0.739	0.216	0.739	0.216	0.739	0.216	0.739	0.216	0.739	0.216
	5% nivel	0.463	0.146	0.463	0.146	0.463	0.146	0.463	0.146	0.463	0.146
	10% nivel	0.347	0.119	0.347	0.119	0.347	0.119	0.347	0.119	0.347	0.119

* Las celdas sombreadas indican el rechazo de la hipótesis nula de raíz unitaria, confirmando la estacionariedad de las series temporales. **Fuente:** elaboración propia.

En cuanto a la determinación de la longitud de los rezagos, se emplearon 3 rezagos, a pesar de que los criterios FPE y HQ sugirieron 1 rezago (Tabla 3). La decisión de optar por 3 rezagos se basó en la ausencia de problemas de correlación serial o inestabilidad en el modelo (Tabla 5).

Tabla 5. Criterio de Longitud de Rezagos

Rez	LogL	LR	FPE	AIC	SC	HQ
0	221.8452	NA	2.8e-12	-12.39115	-12.16896*	-12.31445
1	260.6280	64.26865	1.3e-12*	-13.17874	-11.84559	-12.71854*
2	278.4258	24.40843	2.1e-12	-12.76719	-10.32307	-11.92348
3	297.9419	21.18895	3.8e-12	-12.45382	-8.898743	-11.22661
4	346.6187	38.94142*	1.7e-12	-13.80678	-9.140739	-12.19607
5	380.6641	17.50908	3.5e-12	-14.32366*	-8.546658	-12.32944

* Indica el número de rezagos elegido por cada criterio.

Acronimos por su nombre en inglés: *LR is sequential modified LR test statistic; FPE is Final prediction error; AIC is Akaike information criterion; SC is Schwarz information criterion; HQ is Hannan-Quinn information criterion.* Fuente: elaboración propia.

Ambos modelos VAR cumplen con la condición de estabilidad. Las pruebas de estabilidad demuestran que los módulos son menores a uno en ambos modelos, lo que indica que ninguna raíz se encuentra fuera del círculo unitario (Tabla 6).

Table 6. Prueba de estabilidad.

Modulus
0.792638
0.792638
0.783402
0.783402
0.775646
0.775646

Fuente: elaboración propia.

Verificar la presencia de correlación serial es fundamental, ya que las relaciones estimadas pueden ser espurias si esta existe. Para su detección, se utiliza la prueba LM, cuya hipótesis nula es la existencia de correlación serial en cada rezago. La Tabla 7 muestra que ambos modelos carecen de correlación serial. El modelo no presenta correlación serial después del tercer rezago.

Tabla 7. Prueba LM de correlación.

Rez.	LRE* stat	Prob.
		0.214
1	30.25979	7
		0.199
2	30.67708	9
		0.924
3	15.67483	2
		0.201
4	30.62600	7
		0.444
5	25.32540	3
		0.864
6	17.46783	1

Fuente: Elaboración propia.

En consecuencia, el VAR que se estima tiene a siguiente estructura

$$Y_t = G_0 + G_1 Y_{t-1} + G_2 Y_{t-2} + \dots + G_N Y_{t-14} + e_t, \quad (4)$$

El VAR en la ecuación 1 está compuesto por un vector de 5x1 interceptos, 3 rezagos, y 5 variables endógenas (tasa de interés nominal, PIB, inflación, tipo de cambio real y contagios por COVID). Y_t es un vector de 7x1 variables endógenas, G_0 es un vector de 7x1 interceptos, G_j es una matriz de 7x7 coeficientes para $j = 1, 2, 3$; e_t es un vector 7x1 de innovaciones de ruido blanco.

La identificación de la matriz A se realiza mediante el método recursivo de Sims (1992), por lo que la matriz A es triangular inferior. Esta estructura económica sugiere la existencia de una cadena secuencial de eventos. La identificación de los choques estructurales se justifica teóricamente en Galí y Monacelli (2005). El orden recursivo para el VAR es (Contagios, PIB, tipo de cambio real, inflación, Tasa de interés nominal) Entonces, la matriz identificada es

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ a_{21} & 1 & 0 & 0 & 0 \\ a_{31} & a_{32} & 1 & 0 & 0 \\ a_{41} & a_{42} & a_{43} & 1 & 0 \\ a_{51} & a_{52} & a_{53} & a_{54} & 1 \end{pmatrix}$$

Donde cada $a_{i,j}$ representa las correlaciones contemporáneas entre las variables endógenas. Por lo tanto, la tasa de política monetaria reacciona contemporáneamente a cada uno de los choques estructurales exógenos, lo que también implica que la política monetaria afecta a las variables macroeconómicas un periodo después de que la tasa de interés ha cambiado.

En este análisis, se estiman las Funciones de Impulso Respuesta para doce periodos, considerando respuestas a innovaciones de una desviación estándar de Cholesky dentro de un rango de más/menos 2 errores estándar. Los hallazgos se presentan en el mismo orden que el canal de transmisión teórico de la política monetaria: primero la Regla de Taylor, seguida de la función IS, y luego la Curva de Phillips (Figura 3). Estos resultados se discutirán en la sección correspondiente.

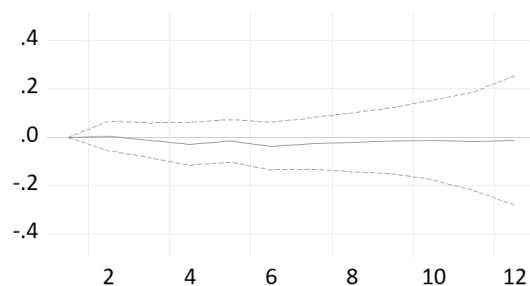
La tasa de interés nominal muestra insensibilidad a los cambios en el PIB. En respuesta al aumento de la inflación, la tasa de interés nominal aumenta gradualmente durante doce meses, alcanzando un máximo de 0.062% en el cuarto mes, para luego disminuir a 0.018% en el duodécimo mes. De manera similar, ante un aumento del tipo de cambio real, la tasa de interés nominal presenta una tendencia al alza, con respuestas máximas de 0.030% y 0.033% en el tercer y sexto mes, respectivamente, disminuyendo gradualmente a 0.003% en el duodécimo mes. En la sección de discusión se explicará cómo el banco central inicialmente priorizó la estabilidad del PIB, pero el aumento de la inflación llevó a un cambio de enfoque hacia su contención (Figura 2).

Con respecto a la función IS, las estimaciones durante el periodo de la pandemia sugieren que la relación teórica se mantiene. Tras un aumento en la tasa de interés nominal, la demanda agregada disminuye un 0.008% en el segundo mes, seguido de una atenuación de su respuesta. Como se discutirá más adelante, esta diferencia se atribuye a variaciones en las magnitudes de las fluctuaciones de la tasa de interés entre los dos periodos (Figura 3).

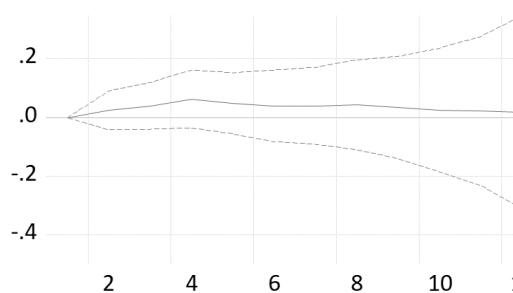
En cuanto a la Curva de Phillips, las estimaciones difieren de la teoría, sin evidencia de una respuesta inflacionaria positiva ante un aumento de la producción. Como se detallará en la sección de discusión, esta diferencia puede atribuirse a la magnitud de los choques de costos experimentados durante la pandemia, los cuales han llevado a una inflación elevada y persistente a nivel global (Figura 4).

Figura 2. FIR: Respuestas de la tasa de interés nominal ante un aumento del PIB, la tasa de inflación y el tipo de cambio real (respuestas a innovaciones de una desviación estándar de Cholesky con un rango de más/menos 2 errores estándar).

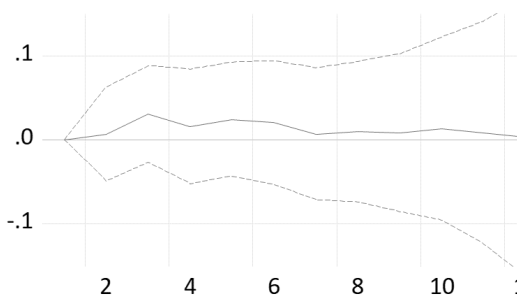
a) Respuestas de la tasa de interés nominal ante un aumento del PIB



b) Respuestas de la tasa de interés nominal ante un aumento de la tasa de inflación

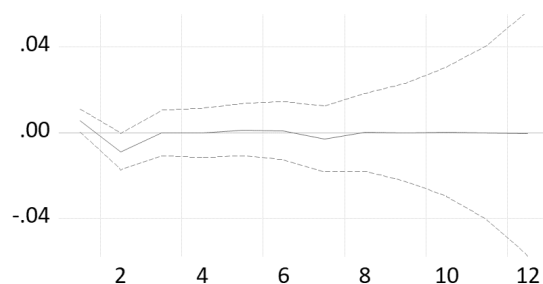


c) Respuestas de la tasa de interés nominal ante un aumento del tipo de cambio real



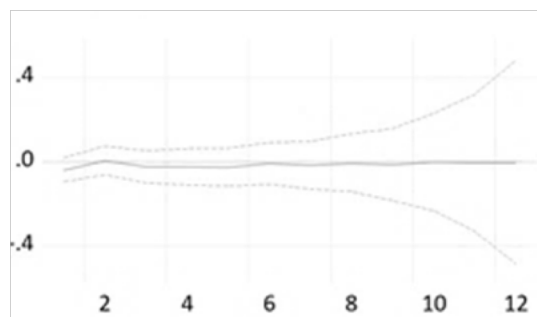
Fuente: elaboración propia.

Figura 3. FIR: Respuesta del PIB ante un aumento de la tasa de interés nominal (respuestas a innovaciones de una desviación estándar de Cholesky con un rango de más/menos 2 errores estándar)..



Fuente: elaboración propia.

Figura 4. FIR: Respuesta de la tasa de inflación ante un aumento del PIB (respuestas a innovaciones de una desviación estándar de Cholesky con un rango de más/menos 2 errores estándar).



Fuente: elaboración propia.

En síntesis, durante la pandemia, la política monetaria centra su enfoque exclusivamente en la tasa de inflación y el tipo de cambio. Es notable que, en el periodo de estabilidad, el PIB no responde a las fluctuaciones de la tasa de interés, pero este patrón cambia durante la pandemia. Mientras que la inflación muestra una respuesta positiva ante incrementos en la producción durante la estabilidad, tal respuesta está ausente durante la pandemia.

Como se mostrará, a pesar de las variaciones en las estimaciones, la combinación de estas con los hallazgos de la literatura indica una transmisión coherente de la política monetaria.

DISCUSIÓN

Esta sección compara las estimaciones del estudio con la literatura existente, abordando cada componente del canal de transmisión de la política monetaria.

En las estimaciones de este artículo, la tasa de interés nominal no reacciona a los cambios en el PIB durante la pandemia. Rodríguez et al. (2023) explican que, al inicio de la pandemia, no fue necesario estabilizar la inflación, ya que hubo choques de oferta positivos y negativos que se contrarrestaron mutuamente. Por lo tanto, el banco central redujo la tasa de interés para evitar una mayor caída del PIB; sin embargo, la depreciación del tipo de cambio y sus presiones inflacionarias incentivaron al Banco de México a aumentar la tasa de interés, a pesar de que la actividad económica seguía siendo baja. Un patrón similar fue observado por Ros (2015), quien cita un estudio de la OCDE que examina el periodo de 2001 a 2008 con datos mensuales y reporta que la tasa de interés nominal busca la estabilidad del PIB solo en situaciones de caídas abruptas, como la caída del 15% registrada durante la pandemia en abril de 2020, después de lo cual la tasa de interés nominal comenzó a disminuir.

Las estimaciones muestran que, en respuesta al aumento de la inflación, la tasa de interés nominal aumenta. Como se mencionó anteriormente, a pesar de que la actividad económica no se recuperó a los niveles previos a la pandemia, la tasa de interés nominal se enfocó nuevamente en estabilizar la tasa de inflación a mediados de 2021. La tasa de política comenzó a aumentar en julio de 2021 (4.3%), alcanzando el 11.15% en junio de 2023. Esto se debió a que la tasa de inflación comenzó a aumentar constantemente en 2021, un incremento que se detendría hacia finales de 2022 debido a múltiples acciones sobre la demanda y la oferta, como la sustitución del consumo de servicios por bienes, el aumento de la demanda agregada de las economías avanzadas que se recuperaron más rápidamente que las emergentes, la escasez de ciertos insumos a medida que los países se recuperaban de la pandemia a un ritmo diferente, lo que se vio exacerbado por la guerra en Ucrania. Así, el banco central aumentó constantemente la tasa de interés nominal, enfrentando choques de oferta pasados que no habían desaparecido y nuevos choques, todos ellos ejerciendo presión sobre la inflación (Rodríguez et al., 2023).

El banco central no solo intervenía para estabilizar la inflación observada, sino que también enfocaba sus esfuerzos en anclar las expectativas de inflación (Rodríguez et al., 2023). La razón de esto era preservar la credibilidad de los agentes económicos respecto a las acciones del banco central para estabilizar la inflación. Para explicar este fenómeno, es importante mencionar la crisis de 1995, cuando la inflación aumentó desanclando las expectativas de inflación. Después de este episodio, las expectativas retrospectivas desempeñaron un papel significativo en el establecimiento de la inflación observada, ya que

los agentes económicos anticiparon que la inflación futura sería alta, como en el pasado. A medida que el banco central intervino de manera consistente y creíble siguiendo su mandato legal —procurar el poder adquisitivo del peso—, disminuyó la influencia de las expectativas retrospectivas en la inflación observada, lo que facilitó la estabilización de los precios futuros (Capistrán y Ramos, 2006; Ros 2015; Sidaoui y Ramos, 2008).

En las estimaciones, la tasa de interés nominal muestra una respuesta a los cambios en el tipo de cambio real. Específicamente, los resultados indican que cuando el tipo de cambio real se deprecia, la tasa de interés nominal aumenta, con el objetivo de apreciarlo y mitigar así la presión sobre la inflación interna, alineándose con las expectativas teóricas (Galí y Monacelli, 2005). Sin embargo, la evidencia empírica para el contexto mexicano introduce complejidad en esta relación, ya que el Banco de México exhibe reacciones asimétricas ante las fluctuaciones del tipo de cambio real: cuando el tipo de cambio real se deprecia, la tasa de interés nominal aumenta, resultando en la apreciación del tipo de cambio real; por el contrario, cuando el tipo de cambio real se aprecia, la tasa de interés nominal permanece sin cambios, en contra de las expectativas teóricas de una disminución que induciría a la depreciación.

Esta divergencia se atribuye a la decisión estratégica del Banco Central de permitir que el tipo de cambio real se aprecie, mitigando la presión inflacionaria asociada con la depreciación (Cermeño et al., 2012). El banco central dio una importancia significativa a mantener la estabilidad del tipo de cambio real, al punto de que limitó su intervención para estimular la actividad económica a principios de 2020 (Rodríguez et al., 2023).

Las estimaciones indican que la demanda agregada responde a los aumentos en la tasa de interés nominal durante la pandemia. Este es un resultado notable dado que la demanda agregada en México típicamente no responde a los cambios en la tasa de política. Aunque no es completamente novedoso, este hallazgo confirma que aumentos sustanciales en la tasa de interés nominal pueden influir en el comportamiento de los agentes económicos en México.

Sidaoui y Ramos (2008), Carvalho y Moura (2010) y Ros (2015) reconocen que la demanda agregada no es completamente insensible a los cambios en la tasa de interés nominal, sino que es inelástica, por lo que se requieren aumentos sustanciales en las tasas de interés para una reacción notable en la demanda agregada. El banco central aumentó consistentemente la tasa de interés nominal para contrarrestar tanto los choques de costos previos como los nuevos durante la pandemia. Incrementos sucesivos de 50 y 70 puntos base, superando las normas históricas recientes, llevaron a un aumento total de 700 puntos base en un año y medio, desde junio de 2021 hasta diciembre de 2022 (Rodríguez et al., 2023).

Además, la preferencia mencionada del Banco de México por un tipo de cambio real apreciado tanto durante la pandemia como en los periodos de estabilidad es otro canal a través del cual la tasa de interés afecta la demanda agregada, ya que resulta en bienes producidos localmente más caros en comparación con los bienes producidos en el extranjero, disminuyendo así la demanda de bienes producidos localmente, tanto dentro como fuera del país (Ros, 2015).

Las estimaciones revelan que la tasa de inflación no responde a los aumentos en la tasa de interés nominal durante la pandemia. Esto puede deberse a que la inflación alta y persistente a nivel global continúa, lo que complica la distinción de los cambios inducidos en la producción sobre la tasa de inflación. Varios choques de costos positivos durante la pandemia, como el aumento de la demanda externa de los países de altos ingresos, la escasez de insumos debido a la reapertura no sincronizada de las economías, la lenta recuperación de la oferta agregada en comparación con la demanda, y la volatilidad de precios inducida por el conflicto en Ucrania, exacerban este desafío (Kashkari, 2024; IMF, 2023 a,b,c; Rodríguez et al., 2023).

Aunque no se explora en este estudio, el canal de expectativas subraya la dependencia del banco central en aumentar la tasa de interés nominal para anclar las expectativas de inflación y controlar la inflación observada. La anticipación pública de aumentos en la tasa de interés en respuesta al aumento de la inflación fomenta la confianza en la capacidad del banco central para controlar efectivamente la inflación, lo que contribuye a la disminución de las expectativas de inflación. Este proceso se complementa reforzando la credibilidad del banco central a través de la comunicación pública sobre sus acciones, como la publicación de pronósticos de inflación y la divulgación de los votos individuales de cada miembro de la autoridad del banco central (Ros, 2015; Rodríguez et al., 2023).

CONCLUSIONES

La principal conclusión de este estudio es que la tasa de interés es una herramienta fundamental para el banco central en su política monetaria. Con el mandato legal de salvaguardar el poder adquisitivo del peso, la tasa de interés se utiliza para asegurar la estabilidad de la inflación y promover la apreciación del tipo de cambio. Aunque su enfoque principal es el control de la inflación, desde la implementación del régimen de objetivos de inflación, la tasa de interés también ha jugado un papel en la estabilización de la actividad económica durante periodos de estabilidad inflacionaria combinados con caídas en el PIB.

La tasa de interés nominal influye en la demanda agregada, aunque esta no responde fácilmente a cambios en la tasa de interés. Se requieren ajustes sustanciales en la tasa de interés nominal para inducir una

reacción en la demanda agregada. Sin embargo, a través de la tendencia a mantener un tipo de cambio real apreciado, el tipo de cambio restringe la actividad económica al reducir el consumo de bienes producidos localmente.

Además, la tasa de interés puede impactar la inflación a través de dos canales: la producción y las expectativas de inflación. Dado que la tasa de interés tiene una influencia limitada sobre la demanda agregada, el canal de producción se ve restringido. Por lo tanto, las expectativas se convierten en el medio principal para controlar los precios. Después de la crisis de 1995, mediante intervenciones consistentes con su mandato legal y una comunicación transparente con el público, el banco central ganó credibilidad en la estabilización de la inflación. Esto resultó en expectativas de inflación ancladas, lo que condujo a la transición de un marco de objetivos de inflación a uno de objetivos de expectativas de inflación. Consecuentemente, además de gestionar la inflación observada, el banco central se enfoca en mantener la confianza pública a través de intervenciones transparentes para controlar la inflación.

Comprender las complejidades de la política monetaria en México, particularmente en relación con la tasa de interés, es significativo para la población general por varias razones. Primero, la estabilidad de precios, influenciada por la política monetaria, impacta directamente el poder adquisitivo de las personas, asegurando la constancia relativa del valor de la moneda en el tiempo. Durante crisis económicas, como la reciente pandemia, las intervenciones en la tasa de interés pueden emplearse para estabilizar la actividad económica, mitigando así los efectos adversos sobre el empleo y los ingresos. La política monetaria también moldea la demanda agregada, influyendo en las decisiones de gasto de los individuos y, en consecuencia, en la economía en general.

El tipo de cambio, otro aspecto influenciado por la política monetaria tiene implicaciones para aquellos que realizan transacciones internacionales o consumen bienes importados. Un tipo de cambio más estable puede beneficiar a los consumidores al evitar fluctuaciones abruptas en los precios de los productos importados. Además, la transparencia y efectividad de la política monetaria influyen en la confianza de los ciudadanos en la economía. Cuando el banco central es percibido como creíble en el mantenimiento de la estabilidad de precios, las expectativas de inflación se mantienen controladas, fomentando un entorno económico más predecible y permitiendo a los individuos tomar decisiones financieras con mayor confianza.

Entre las principales limitaciones, cabe destacar que las expectativas de inflación no fueron incluidas en los modelos, lo que podría sesgar las estimaciones debido a variables omitidas; investigaciones futuras podrían incorporarlas para proporcionar mayor claridad sobre el canal de transmisión de la política

monetaria. Además, el análisis del periodo desde el primer contagio de Covid-19 en México en febrero de 2020 representa una muestra pequeña, por lo que se considera recomendable revisar el tema de los cambios en el canal de transmisión de la política monetaria cuando se disponga de más datos.

REFERENCIAS

- Ambler, S., y Kronick, J. (2021). For the Record: Assessing the Monetary Policy Stance of the Bank of Canada. CD Howe Institute Commentary, 588. https://www.cdhowe.org/sites/default/files/2022-04/Commentary_588_0.pdf
- Azad, F., Serletis, A., y Xu, L. (2021). Covid-19 and monetary–fiscal policy interactions in Canada. The Quarterly Review of Economics and Finance, 81, 376-384. DOI: 10.1016/j.qref.2021.06.009
- Banco de México (2020). Regional economies report January-March 2020. <https://www.banxico.org.mx/publicaciones-y-prensa/reportes-sobre-las-economias-regionales/reportes-economias-regionales.html>
- Capistrán, C., y Ramos-Francia, M. (2006). Inflation dynamics in Latin America, documento de investigación, núm. 2006-11. Banco de México. <https://core.ac.uk/download/pdf/6278369.pdf>
- Carvalho, A., y Moura, M. L. (2009). What can Taylor rules say about monetary policy in Latin America? Journal of Macroeconomics, 32(1), 392–404. https://econpapers.repec.org/article/eeejmacro/v_3a32_3ay_3a2010_3ai_3a1_3ap_3a392-404.htm
- Cermeño, R., Villagómez, F., y Orellana, J. (2012). Monetary policy rules in a small open economy: An application to Mexico. “Journal of Applied Economics”, 15 (2), pp. 259-286. [https://doi.org/10.1016/S1514-0326\(12\)60012-9](https://doi.org/10.1016/S1514-0326(12)60012-9)
- Clarida, R., Gali, J., y Gertler, M. (1998). Monetary policy rules in practice: Some international evidence. European economic review, 42(6), 1033-1067. [https://doi.org/10.1016/S0014-2921\(98\)00016-6](https://doi.org/10.1016/S0014-2921(98)00016-6)
- Costa, R. (2023). Monetary policy in Chile: combining theory, evidence and experience. Central banking in the Americas: Lessons from two decades, 69. https://www.bis.org/publ/bppdf/bispap143_f.pdf
- Dallari, P., y Ribba, A. (2020). The dynamic effects of monetary policy and government spending shocks on unemployment in the peripheral Euro area countries. Economic Modelling, 85, 218-232. DOI: 10.1016/j.econmod.2019.05.018
- Das, P. (2019) Econometrics in Theory and Practice Analysis of Cross Section, Time Series and Panel Data with Stata 15.1. Springer.

- Eichenbaum, M. S., Rebelo, S., y Trabandt, M. (2021). The macroeconomics of epidemics. *The Review of Financial Studies*, 34(11), 5149-5187. <https://doi.org/10.1093/rfs/hhab040>
- Galí, J., y Monacelli, T. (2005). Monetary policy and exchange rate volatility in a small open economy. *The Review of Economic Studies*, 72(3), 707–734. <https://crei.cat/wp-content/uploads/users/pages/roes8739.pdf>
- Galí, J. (2015). *Monetary policy, inflation, and the business cycle: an introduction to the new Keynesian framework and its applications*. Princeton University Press.
- Guerrieri, V., Lorenzoni, G., Straub, L., y Werning, I. (2022). Macroeconomic implications of COVID-19: Can negative supply shocks cause demand shortages?. *American Economic Review*, 112(5), 1437-1474. DOI: 10.1257/aer.20201063
- Hevia, C., y J. P. Nicolini (2013). Optimal devaluations. *IMF Economic Review*, 61(1), 22-51. <https://ideas.repec.org/a/pal/imfecr/v61y2013i1p22-51.html>
- International Monetary Fund (IMF) (2023a). IMF Country Report No.23/273. <https://www.imf.org/-/media/Files/Publications/CR/2023/English/IITAEA2023001.ashx>
- International Monetary Fund (IMF) (2023b). IMF Country Report No.23/189. <https://www.imf.org/-/media/Files/Publications/CR/2023/English/IPOLEA2023001.ashx>
- International Monetary Fund (IMF) (2023c). IMF Country Report No.23/33. <https://www.imf.org/-/media/Files/Publications/CR/2023/English/IESPEA2023001.ashx>
- Kashkari, N. (2024). Policy Has Tightened a Lot. How Tight Is It? Federal Reserve Bank of Minneapolis. <https://www.minneapolisfed.org/article/2024/policy-has-tightened-a-lot-how-tight-is-it>
- León, I. E. T., y Martínez, F. V. (2013). Principales determinantes en las decisiones de política monetaria de México: un análisis econométrico. *Estudios económicos*, 79-107. <https://www.redalyc.org/pdf/597/59727431003.pdf>
- Loria, E., y Ramírez, J. (2011). Inflation, Monetary Policy and Economic Growth in Mexico. *An Inverse Causation*, 1970-2009. “Scientific Research Publishing”, 2(5), pp. 834-845. <https://www.scirp.org/html/8704.html>
- Ouliaris, S., Pagan, A., y Restrepo, J. (2016) *Quantitative Macroeconomic Modeling with Structural Vector Autoregressions—An EViews Implementation*. I, 238 pp., E-views publishing.
- Rodríguez, V., Cuadra, G., y Sámano, D. (2023). Inflation targeting in Mexico: evolution, achievements and policy lessons. *Central Banking in the Americas: Lessons from Two Decades*, 113. https://www.bis.org/publ/bppdf/bispap143_h.pdf

- Ros, J. (2015). Central bank policies in Mexico: Targets, instruments, and performance. *Comparative Economic Studies*, 57(3), 483-510. <https://ideas.repec.org/a/pal/compes/v57y2015i3p483-510.html>
- Sidaoui, J., y Ramos-Francia, M. (2008), The Monetary Transmission Mechanism in Mexico: Recent Developments, “Transmission Mechanisms for Monetary Policy in Emerging Market Economies”, BIS Papers 35, pp. 363-394. <https://econpapers.repec.org/bookchap/bisbisbpc/35-17.htm>
- Sims, C. (1992). Interpreting the macroeconomic time series facts: The effects of monetary policy. *European Economic Review*, 36(5), 975-1000. [https://doi.org/10.1016/0014-2921\(92\)90041-T](https://doi.org/10.1016/0014-2921(92)90041-T)
- Stock, J. W., y Watson, M. W. (2020). *Introduction to Econometrics—Global Edition* (updated 4th ed.). Harlow: Pearson Education.
- Woodford, M. (2007). Globalization and monetary control, No. w13329. National Bureau of Economic Research.

BASES DE DATOS

- Banco de México (2023a). Valores Gubernamentales. <https://www.banxico.org.mx/SieInternet/consultarDirectorioInternetAction.do?accion=consultarCuadro&idCuadro=CF107§or=22&locale=es>
- Banco de México (2023b). Índice del tipo de cambio real del peso mexicano con precios consumidor, respecto a 49 países ponderados con comercio. <https://www.banxico.org.mx/SieInternet/consultarDirectorioInternetAction.do?accion=consultarCuadro&idCuadro=CR184§or=2&locale=es>
- CONAHCYT (2023). Covid-19 Mexico. Descargas: Casos diarios por Estado + Nacional. <https://datos.covid-19.conacyt.mx/#DownZCSV>
- Fondo Monetario Internacional (IMF) (2023). Data Mapper. https://www.imf.org/external/datamapper/NGDP_RPCH@WEO/OEMDC/ADVEC/WEOWORLD
- INEGI (2023a). Indicador Global de Actividad Económica. <https://www.inegi.org.mx/temas/igae/#Tabulados>
- INEGI (2023b). Banco de Información Económica. <https://www.inegi.org.mx/app/indicadores/?tm=0>

ANÁLISIS DE LA CANTIDAD DE ACCIDENTES DE TRÁNSITO MENSUALES EN EL ESTADO DE PUEBLA MEDIANTE EL MODELO NORMAL UNIVARIADO CON PUNTOS DE CAMBIO

Blanca Xochilt Muñoz Vargas, Bulmaro Juárez Hernández, Hortensia Josefina Reyes Cervantes

Facultad de Ciencias Físico Matemáticas

Benemérita Universidad Autónoma de Puebla

e-mail: b.xochilt.munoz@gmail.com, bjuarez@fcfm.buap.mx, hreyes@fcfm.buap.mx

RESUMEN

El problema del punto de cambio consiste en detectar si existe algún cambio en una sucesión de variables aleatorias observadas y estimar el número de cambios y sus correspondientes ubicaciones. En este trabajo se abordan los problemas del punto de cambio en el modelo Normal univariado: cambio en la media, cambio en la varianza y cambio en la media y la varianza. Estos modelos se aplican a una base de datos de Accidentes de tránsito terrestre en zonas urbanas y suburbanas que han ocurrido en el Estado de Puebla entre 1997 y 2017. El objetivo de este análisis es detectar fechas en las que se tuvo un aumento o una disminución en la cantidad de accidentes de tránsito.

Palabras claves: Punto de cambio, distribución normal, accidentes de tránsito terrestre.

INTRODUCCIÓN

El mundo está lleno de cambios, el conocimiento de esos cambios puede ayudar a las personas a evitar pérdidas innecesarias y aprovechar las transiciones beneficiosas. En muchas situaciones prácticas, el investigador se enfrenta con el problema de detectar el número de puntos de cambio o saltos y sus ubicaciones. Esto se conoce como el *problema del punto de cambio* (Chen y Gupta, 2012).

Usualmente la inferencia estadística sobre los puntos de cambio tiene dos aspectos:

1. Detectar si existe algún cambio en la sucesión de variables aleatorias observadas.
2. Estimar el número de cambios y sus correspondientes ubicaciones (Chen y Gupta, 2012).

Sea $X_1, X_2, \dots, X_n$ una sucesión de variables aleatorias independientes con funciones de distribución de probabilidad $F_1(\cdot), F_2(\cdot), \dots, F_n(\cdot)$, respectivamente. Si las distribuciones $F_1(\cdot), F_2(\cdot), \dots, F_n(\cdot)$ pertenecen

a una familia paramétrica común, $F(\cdot, \underline{\theta})$, donde $\underline{\theta} \in \mathbb{R}^p$, entonces el problema del punto de cambio es contrastar la hipótesis nula sobre los parámetros poblacionales $\underline{\theta}_i$, $i = 1, \dots, n$:

$$H_0: \underline{\theta}_1 = \underline{\theta}_2 = \dots = \underline{\theta}_n = \underline{\theta} \text{ (desconocido)}, \quad (1)$$

contra la hipótesis alternativa

$$H_1: \underline{\theta}_1 = \dots = \underline{\theta}_{k_1} \neq \underline{\theta}_{k_1+1} = \dots = \underline{\theta}_{k_2} \neq \dots \neq \underline{\theta}_{k_{(q-1)}+1} = \dots = \underline{\theta}_{k_q} \neq \underline{\theta}_{k_q+1} = \dots = \underline{\theta}_n, \quad (2)$$

donde $1 < k_1 < \dots < k_{(q-1)} < k_q < n$, q es el número desconocido de puntos de cambio y $k_1, \dots, k_{(q-1)}, k_q$ son las respectivas posiciones desconocidas que se tienen que estimar. Estas hipótesis, juntas, revelan los aspectos de la inferencia del punto de cambio: determinar si existe algún punto de cambio en el proceso y estimar el número y posición (posiciones) del punto (de los puntos) de cambio (Chen y Gupta, 2012).

Vostrikova (1981) propuso un método para detectar el número de puntos de cambio y sus ubicaciones en un proceso aleatorio multidimensional, conocido como el *procedimiento de segmentación binaria*. Utilizando este procedimiento es suficiente trabajar de manera recursiva con el problema de detectar un punto de cambio y su ubicación.

EL PROBLEMA DEL PUNTO DE CAMBIO EN EL MODELO NORMAL UNIVARIADO

El problema de probar y estimar los puntos de cambio ha atraído mucha atención en la literatura, la mayoría de los modelos propuestos para los problemas de punto de cambio son modelos normales (univariados o multivariados), esto se debe a que el modelo normal es el modelo más común en la práctica. El problema del punto de cambio también se ha estudiado en los modelos: exponencial, gamma, de regresión y en supervivencia, por mencionar algunos. En el modelo normal univariado se puede formular el problema del punto de cambio de diferentes formas:

- punto de cambio en la media,
- punto de cambio en la varianza,
- punto de cambio en la media y la varianza.

Punto de cambio en la media

Suponga que cada X_i esta distribuida normalmente con media μ_i y varianza común σ^2 , $i = 1, 2, \dots, n$. En el problema del punto de cambio en la media la hipótesis de estabilidad (hipótesis nula) se define como

$$H_0: \mu_1 = \mu_2 = \dots = \mu_n = \mu. \quad (3)$$

El interés es contrastar (3) contra

$$H_1: \mu_1 = \dots = \mu_k \neq \mu_{k+1} = \dots = \mu_n, \quad (4)$$

donde k es la ubicación desconocida del único punto de cambio. El procedimiento de prueba depende de si el parámetro de ruido σ^2 es conocido o desconocido.

Punto de cambio en la media con varianza conocida.

Sin pérdida de generalidad, se supone que $\sigma^2 = 1$. Bajo H_0 , la función de verosimilitud es

$$L_0(\mu) = \frac{1}{(\sqrt{2\pi})^n} e^{-\frac{1}{2}\sum_{i=1}^n (X_i - \mu)^2}$$

y el estimador de máxima verosimilitud para μ es

$$\hat{\mu} = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i.$$

Bajo H_1 , la función de verosimilitud es

$$L_1(\mu_1, \mu_n) = \frac{e^{-\frac{1}{2}(\sum_{i=1}^k (X_i - \mu_1)^2 + \sum_{i=k+1}^n (X_i - \mu_n)^2)}}{(\sqrt{2\pi})^n}$$

y los estimadores de máxima verosimilitud para μ_1 y μ_n son

$$\hat{\mu}_1 = \bar{X}_k = \frac{1}{k} \sum_{i=1}^k X_i \quad \text{y} \quad \hat{\mu}_n = \bar{X}_{n-k} = \frac{1}{n-k} \sum_{i=k+1}^n X_i.$$

Sea

$$T_k = \sqrt{\frac{n}{k(n-k)}} \left[\sum_{i=1}^k (X_i - \bar{X}) \right],$$

entonces

$$U = \max_{1 \leq k \leq n-1} |T_k|,$$

es un estadístico de prueba basado en la verosimilitud para contrastar H_0 (3) contra H_1 (4).

Hawkins (1977) derivó la distribución nula exacta del estadístico de prueba U , la cual se enuncia en el siguiente teorema.

Teorema 1. La función de densidad de probabilidad nula de U esta dada por

$$f_U(x) = 2 \phi(x, 0, 1) \sum_{k=1}^{n-1} g_k(x, x) g_{n-k}(x, x),$$

donde $\phi(\cdot, 0, 1)$ es la función de densidad de probabilidad de una $N(0, 1)$, $g_1(x, s) = 1$ para $x, s \geq 0$, y $g_k(x, s) = P\{|T_i| < s, i = 1, \dots, k-1 | |T_k| = x\}$, para $x, s \geq 0$.

La demostración de este teorema se puede consultar en Hawkins (1977) y en Chen y Gupta (2012).

Aunque se ha obtenido la distribución nula del estadístico de prueba U , la evaluación de $g_k(\cdot, \cdot)$ requiere cálculos moderados, por lo que Yao y Davis (1986) han derivado la distribución nula asintótica de U , lo cual proporciona una forma alternativa para hacer análisis estadístico cuando n es suficientemente grande, tal distribución está dada en el siguiente teorema, el cual indica que la distribución nula asintótica de U es una distribución Gumbel.

Teorema 2. Bajo H_0 (3), para $-\infty < x < \infty$,

$$\lim_{n \rightarrow \infty} P[a_n^{-1} (U - b_n) \leq x] = \exp\left\{-2\pi^{\frac{1}{2}} - e^{-x}\right\},$$

donde $a_n = (2 \log \log n)^{-\frac{1}{2}}$ y $b_n = a_n^{-1} + \frac{1}{2} a_n \log \log \log n$.

La demostración de este teorema se puede consultar en Yao y Davis (1986) y en Chen y Gupta (2012).

Punto de cambio en la media con varianza desconocida.

Bajo H_0 , la función de verosimilitud es

$$L_0(\mu, \sigma^2) = \frac{1}{(\sqrt{2\pi}\sigma^2)^n} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (X_i - \mu)^2}$$

y los estimadores de máxima verosimilitud para μ y σ^2 son $\hat{\mu} = \bar{X}$ y

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2.$$

Bajo H_1 , la función de verosimilitud es

$$L_1(\mu_1, \mu_n, \sigma_1^2) = \frac{e^{-\frac{1}{2\sigma_1^2} (\sum_{i=1}^k (X_i - \mu_1)^2 + \sum_{i=k+1}^n (X_i - \mu_n)^2)}}{(\sqrt{2\pi}\sigma_1^2)^n}$$

y los estimadores de máxima verosimilitud para μ_1 , μ_n y σ_1^2 son $\hat{\mu}_1 = \bar{X}_k$, $\hat{\mu}_n = \bar{X}_{n-k}$ y

$$\hat{\sigma}_1^2 = \frac{1}{n} \left(\sum_{i=1}^k (X_i - \bar{X}_k)^2 + \sum_{i=k+1}^n (X_i - \bar{X})^2 \right).$$

Sean

$$S^2 = \sum_{i=1}^n (X_i - \bar{X})^2 \quad \text{y} \quad T_k^2 = \frac{k(n-k)}{n} (\bar{X}_k - \bar{X}_{n-k})^2.$$

El estadístico de prueba basado en la razón de verosimilitud entonces está dado por

$$V = \max_{1 \leq k \leq n-1} \frac{|T_k|}{S}.$$

Worsley (1979) obtuvo la distribución nula de V .

Punto de cambio en la varianza.

Sea $X_1, X_2, \dots, X_n$ una sucesión de variables aleatorias normales independientes con parámetros $(\mu, \sigma_1^2), (\mu, \sigma_2^2), \dots, (\mu, \sigma_n^2)$, respectivamente. Se supone que μ es conocida digamos $\mu = \mu_0$. El interés aquí es sobre el cambio en la varianza, por lo que se contrasta la hipótesis de estabilidad (hipótesis nula) definida como

$$H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2 \quad (5)$$

contra la alternativa

$$H_1: \sigma_1^2 = \dots = \sigma_k^2 \neq \sigma_{k+1}^2 = \dots = \sigma_n^2, \quad (6)$$

donde k es la ubicación desconocida del único punto de cambio.

Bajo H_0 , el log de la función de verosimilitud es

$$\log L_0(\sigma^2) = -\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (X_i - \mu_0)^2$$

y el estimador de máxima verosimilitud para σ^2 es

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu_0)^2.$$

Bajo H_1 , el log de la función de verosimilitud es

$$\begin{aligned} \log(L_1(\sigma_1^2, \sigma_n^2)) &= -\frac{n}{2} \log(2\pi) - \frac{k}{2} \log(\sigma_1^2) - \frac{n-k}{2} \log(\sigma_n^2) \\ &\quad - \frac{1}{2\sigma_1^2} \sum_{i=1}^k (X_i - \mu_0)^2 - \frac{1}{2\sigma_n^2} \sum_{i=k+1}^n (X_i - \mu_0)^2 \end{aligned}$$

y los estimadores de máxima verosimilitud para σ_1^2 y σ_n^2 son

$$\hat{\sigma}_1^2 = \frac{1}{k} \sum_{i=1}^k (X_i - \bar{X}_k)^2 \quad \text{y} \quad \hat{\sigma}_n^2 = \frac{1}{n-k} \sum_{i=k+1}^n (X_i - \bar{X}_k)^2.$$

El estadístico de razón de verosimilitud es

$$\lambda_n = \left(\max_{1 \leq k \leq n-1} n \log \hat{\sigma}^2 - k \log \hat{\sigma}_1^2 - (n-k) \log \hat{\sigma}_n^2 \right)^{\frac{1}{2}}. \quad (7)$$

Notar que, para obtener los estimadores de máxima verosimilitud, sólo se pueden detectar cambios para $2 \leq k \leq n-2$. De acuerdo con el principio del procedimiento de máxima verosimilitud, se estima la

posición del punto de cambio, k , con $\hat{k}$ tal que (7) alcanza su máximo en $\hat{k}$. El siguiente teorema proporciona la distribución nula asintótica de λ_n .

Teorema 3. Bajo la hipótesis nula H_0 (5), cuando $n \rightarrow \infty$, $k \rightarrow \infty$, tal que $(k/n) \rightarrow \infty$, entonces para todo $x \in \mathbb{R}$,

$$\lim_{n \rightarrow \infty} P [a(\log n)\lambda_n - b(\log n) \leq x] = \exp\{-2e^{-x}\},$$

donde $a(\log n) = (2 \log \log n)^{1/2}$ y $b(\log n) = 2 \log \log n + \frac{1}{2} \log \log \log n - \Gamma\left(\frac{1}{2}\right)$.

La demostración de este teorema se puede consultar en Chen y Gupta (2012).

Punto de cambio en la media y la varianza.

Sea $X_1, X_2, \dots, X_n$ una sucesión de variables aleatorias normales independientes con parámetros (μ_1, σ_1^2) , (μ_2, σ_2^2) , ..., (μ_n, σ_n^2) , respectivamente. El interés aquí es sobre el cambio en la media y la varianza, por lo que se contrasta la hipótesis de estabilidad (hipótesis nula) definida como

$$\begin{aligned} H_0: \mu_1 = \mu_2 = \dots = \mu_n = \mu. \\ \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2 \text{ (}\mu \text{ y } \sigma^2 \text{ desconocidas),} \end{aligned} \quad (8)$$

contra la alternativa

$$\begin{aligned} H_1: \mu_1 = \dots = \mu_k \neq \mu_{k+1} = \dots = \mu_n \\ \sigma_1^2 = \dots = \sigma_k^2 \neq \sigma_{k+1}^2 = \dots = \sigma_n^2, \end{aligned} \quad (9)$$

donde k es la ubicación desconocida del único punto de cambio. Bajo H_0 , el log de la función de verosimilitud es

$$\log L_0(\mu, \sigma^2) = -\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (X_i - \mu)^2$$

y los estimadores de máxima verosimilitud para μ y σ^2 son $\hat{\mu} = \bar{X}$ y

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2.$$

Bajo H_1 , el log de la función de verosimilitud es

$$\begin{aligned} \log(L_1(\mu_1, \mu_n, \sigma_1^2, \sigma_n^2)) = & -\frac{n}{2} \log(2\pi) - \frac{k}{2} \log(\sigma_1^2) - \frac{n-k}{2} \log(\sigma_n^2) \\ & - \frac{1}{2\sigma_1^2} \sum_{i=1}^k (X_i - \mu)^2 - \frac{1}{2\sigma_n^2} \sum_{i=k+1}^n (X_i - \mu)^2 \end{aligned}$$

y los estimadores de máxima verosimilitud para μ_1 , μ_n , σ_1^2 y σ_n^2 son $\hat{\mu}_1 = \bar{X}_k$, $\hat{\mu}_n = \bar{X}_{n-k}$,

$$\hat{\sigma}_1^2 = \frac{1}{k} \sum_{i=1}^k (X_i - \bar{X}_k)^2 \quad \text{y} \quad \hat{\sigma}_n^2 = \frac{1}{n-k} \sum_{i=k+1}^n (X_i - \bar{X}_{n-k})^2.$$

El estadístico de razón de verosimilitud es

$$\lambda_n = \left(\max_{2 \leq k \leq n-2} n \log \hat{\sigma}^2 - k \log \hat{\sigma}_1^2 - (n-k) \log \hat{\sigma}_{n-k}^2 \right)^{\frac{1}{2}}.$$

Horváth (1993) derivó la distribución nula asintótica de λ_n .

Teorema 4. Bajo la hipótesis nula H_0 (8), cuando $n \rightarrow \infty$,

$$\log_{n \rightarrow \infty} P[a(\log n)\lambda_n - b(\log n) \leq x] = \exp\{-2e^{-x}\}$$

para $x \in \mathbb{R}$, donde $a(\log n) = (2 \log \log n)^{1/2}$ y $b(\log n) = 2 \log \log n + \log \log \log n$. La demostración de este teorema se puede consultar en Horváth (1993) y en Chen y Gupta (2012).

CASO DE ESTUDIO: ANÁLISIS DE LA CANTIDAD DE ACCIDENTES DE TRÁNSITO TERRESTRE EN PUEBLA

Los primeros datos que dan cuenta de los accidentes de tránsito terrestreⁱ en México se remontan a 1928, año en que inicia la estadística con el registro exclusivamente de los atropellamientos. A partir del siguiente año la temática se amplió gradualmente, de tal forma que se comenzó a recopilar información sobre choques, volcaduras de vehículos y caídas de pasajeros, además de los atropellamientos ya captados (INEGI, 2016).

Con el propósito de mejorar la generación de la estadística accidentes de tránsito terrestre en zonas urbanas y suburbana (ATUS), en 1987 se reestructura el cuestionario de captación considerando las recomendaciones internacionales existentes, al mismo tiempo que se inician las pruebas para el procesamiento y generación de resultados por medio de un sistema informático integral (INEGI, 2016).

En 1997 la descentralización del tratamiento de la información sobre accidentes de tránsito permitió aprovechar la estructura operativa del INEGI integrada por diez Direcciones Regionales y 32 Coordinaciones Estatales, con lo que disminuyeron los rezagos en la entrega de datos y se mejoró la oportunidad en su divulgación. Cabe señalar que hasta 1996 el tratamiento de la estadística se realizó en forma centralizada (INEGI, 2016).

Disponer de información sobre la estadística ATUS, permite conocer y cuantificar los accidentes de tránsito terrestre y sus consecuencias para apoyar los estudios de siniestralidad que presenta el tránsito urbano y suburbano de vehículos automotores y de esta forma coadyuvar a la planeación, prevención y organización de la vialidad a nivel regional (INEGI, 2016).

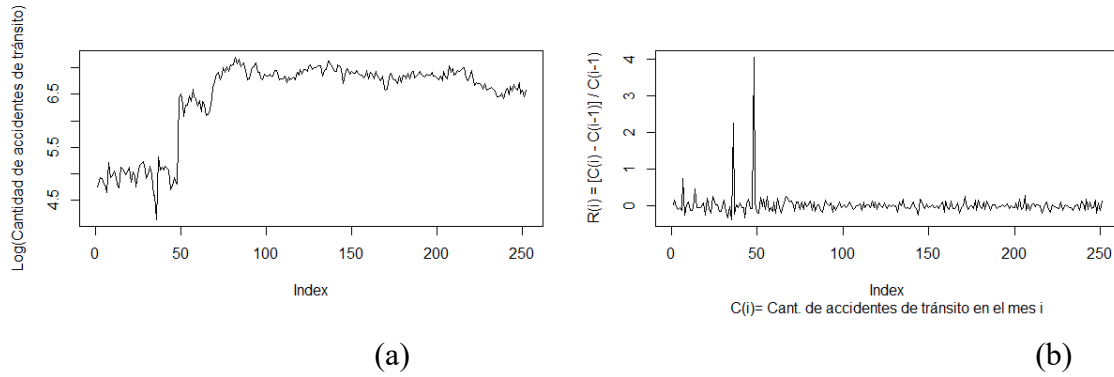


Figura 1. (a) Gráfica de Y_i , $i = 1, \dots, 252$. (b) Gráfica de R_i , $i = 1, \dots, 251$.

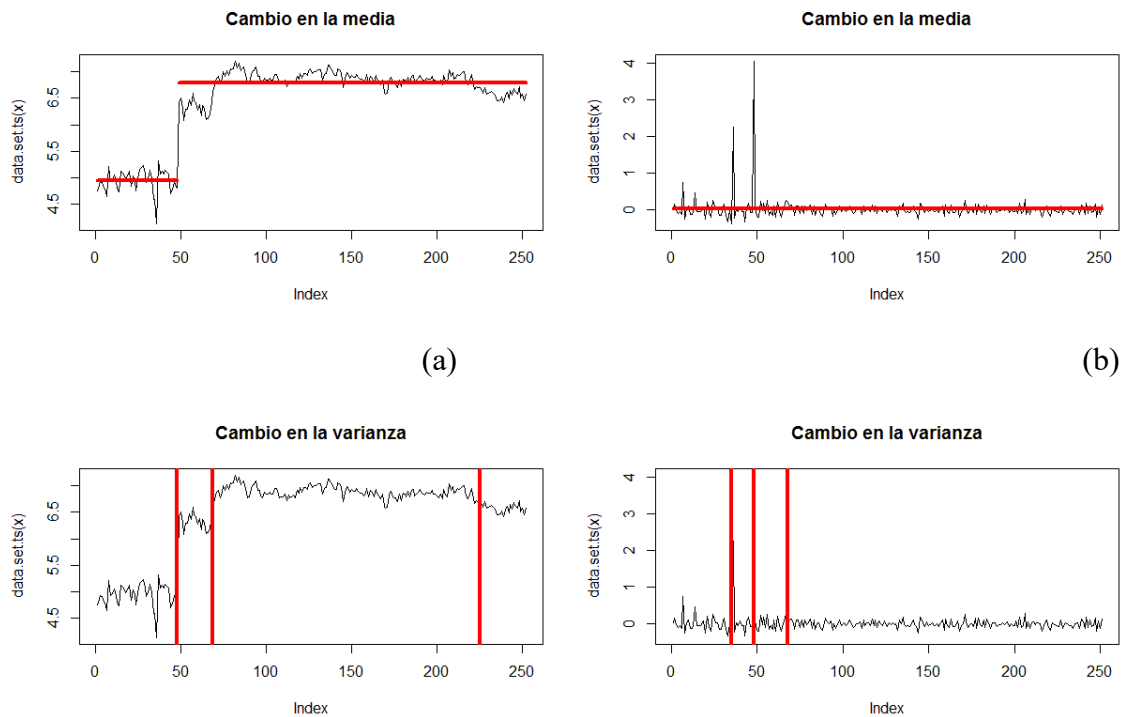
En este trabajo se analiza la cantidad de accidentes de tránsito terrestre que ocurrieron por mes en el Estado de Puebla entre 1997 y 2017 por lo que se tienen 252 datos. Sea X_i = Cant. de accidentes de tránsito terrestres en el mes i , $i = 1, \dots, 252$. Se aplica el modelo Normal univariado con punto de cambio en la media, en la varianza y en la media y la varianza a

$$Y_i = \log X_i, i = 1, \dots, 252$$

y a

$$R_i = \frac{X_{i+1} - X_i}{X_i}, i = 1, \dots, 251.$$

El análisis se hace usando la librería *changepoint* de R.



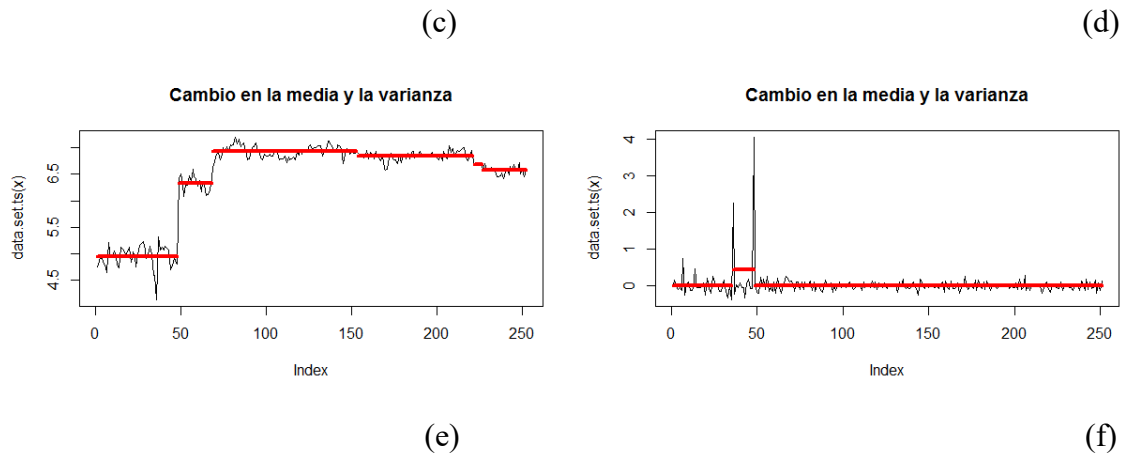


Figura 2. Gráfica de $Y_i, i = 1, \dots, 252$ con punto de cambio en la media (a), en la varianza (c) y en la media y la varianza (e). Gráfica de $R_i, i = 1, \dots, 251$ con punto de cambio en la media (b), en la varianza (d) y en la media y la varianza (f).

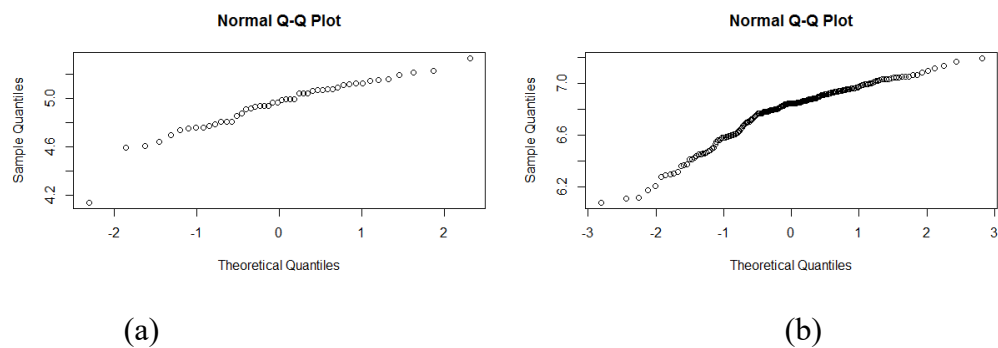
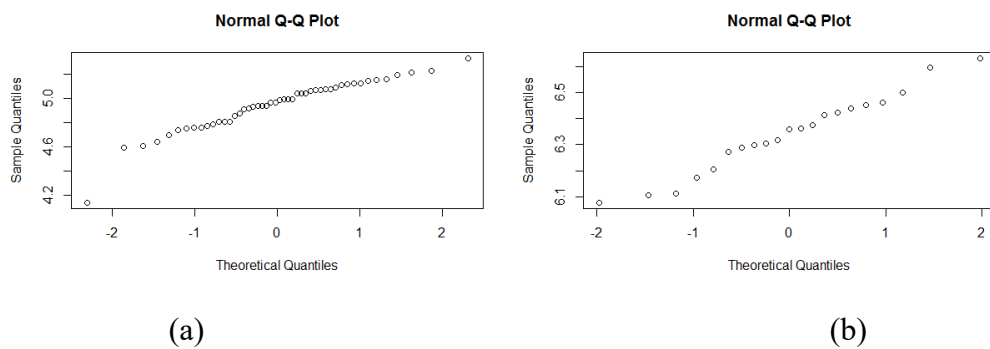
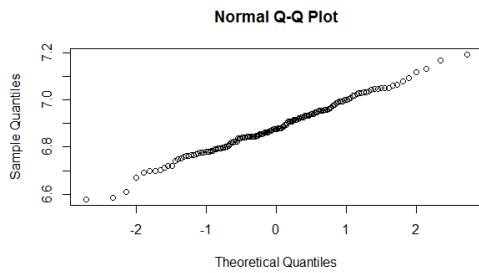
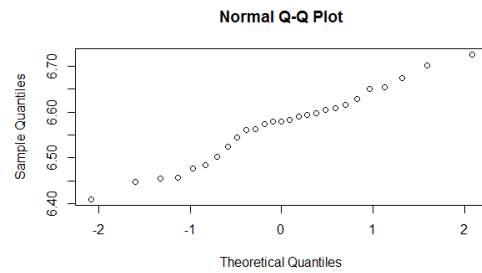


Figura 3. Q-Q plot de Y_i con (a) $i = 1, \dots, 48$ y (b) $i = 49, \dots, 252$.



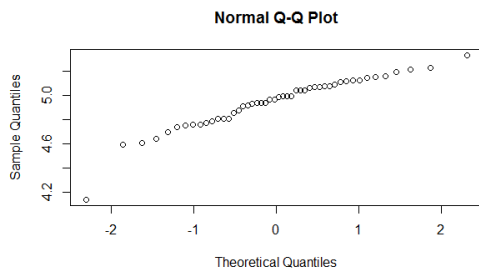


(c)

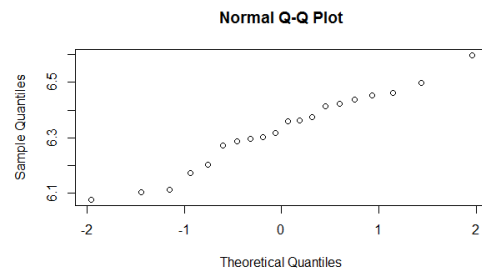


(d)

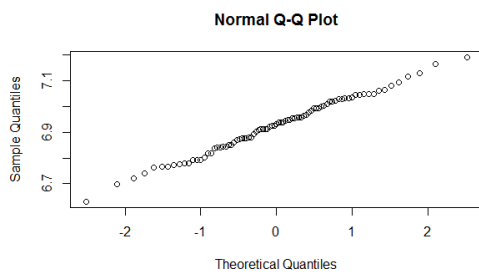
Figura 4. Q-Q plot de Y_i con (a) $i = 1, \dots, 48$, (b) $i = 49, \dots, 69$, (c) $i = 70, \dots, 225$ y (d) $i = 226, \dots, 252$.



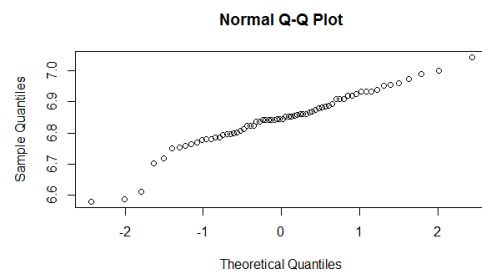
(a)



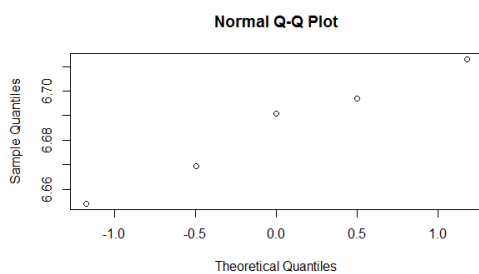
(b)



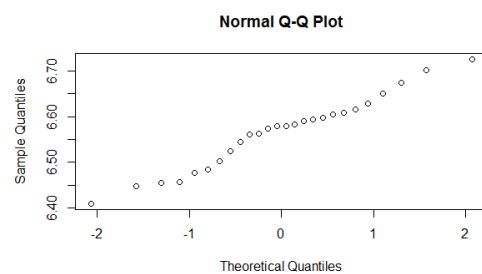
(c)



(d)



(e)



(f)

Figura 5. Q-Q plot de Y_i con (a) $i = 1, \dots, 48$, (b) $i = 49, \dots, 68$, (c) $i = 69, \dots, 153$, (d) $i = 154, \dots, 221$, (e) $i = 222, \dots, 226$ y (f) $i = 227, \dots, 252$.

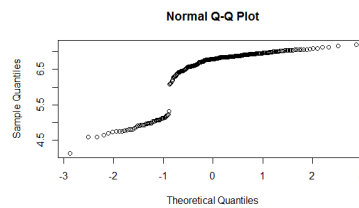


Figura 6. Q-Q plot de R_i , $i = 1, \dots, 251$.

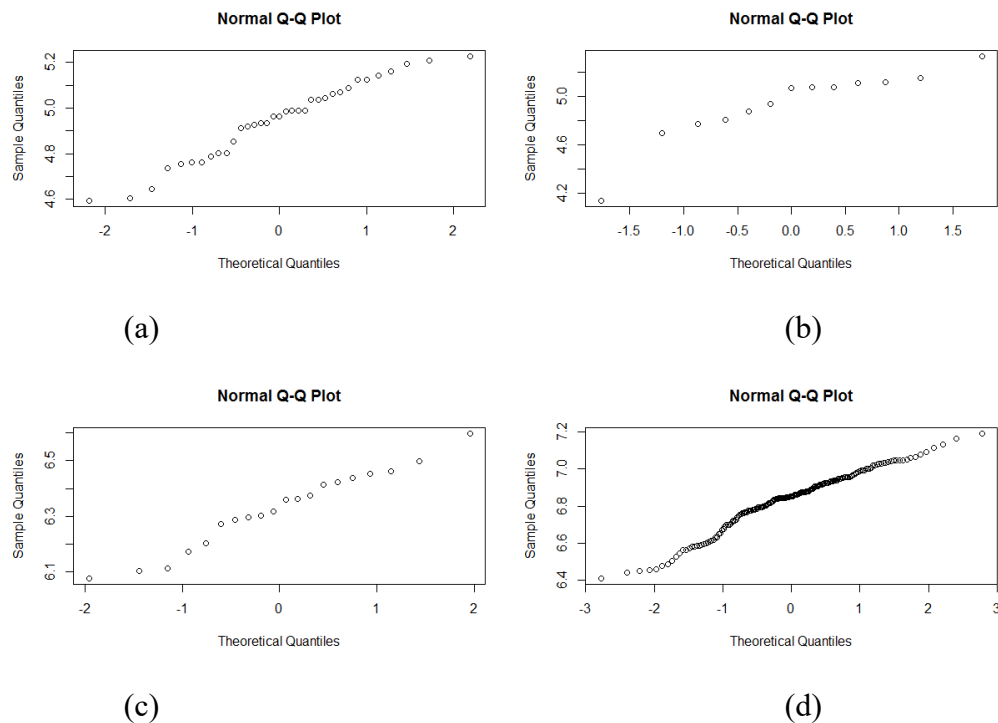


Figura 7. Q-Q plot de R_i con (a) $i = 1, \dots, 35$, (b) $i = 36, \dots, 48$, (c) $i = 49, \dots, 68$ y (d) $i = 69, \dots, 251$.

Tabla 1. Puntos de cambio (PC's) obtenidos con las diferentes metodologías

PC's en	PC's	PC's
la media	48	-
la varianza	48, 69, 225	35, 48, 68
la media y la varianza	48, 68, 153, 221, 226	35, 48

La Figura 1 muestra la gráfica de $Y_i, i = 1, \dots, 252$ y de $R_i, i = 1, \dots, 251$. En la Figura 2 se pueden observar los puntos de cambio que se tienen cuando se supone que $Y_i, i = 1, \dots, 252$ y de $R_i, i = 1, \dots, 251$ tienen una distribución normal con puntos de cambio. En la Tabla 1 se enlistan los puntos de cambio obtenidos con las diferentes metodologías. Se nota que la mayoría de las metodologías en el análisis del punto de cambio detectan un punto de cambio en 48 y en 68, por lo que es de interés investigar que paso en los periodos Diciembre del 2000-Enero de 2001 y en Agosto-Septiembre de 2002. De las gráficas se observa que en estos periodos hubo un aumento en la cantidad de accidentes de tránsito terrestres.

En la Figura 3 se muestra la Q-Q plot de los segmentos de datos que proporciona el análisis con punto de cambio en la media, estos son: (a) $Y_i, i = 1, \dots, 48$, y (b) $Y_i, i = 49, \dots, 252$. De manera análoga en las figuras 4 y 5 se presentan las Q-Q plots de los segmentos de datos que proporciona el análisis con punto de cambio en la varianza (Figura 4) y en la media y la varianza (Figura 5), respectivamente, cuando se analiza $Y_i, i = 1, \dots, 252$. Las Figuras 6, 7 y 8 presentan las Q-Q plots de los segmentos de datos que proporciona el análisis con punto de cambio en la media (Figura 6), en la varianza (Figura 7) y en la media y la varianza (Figura 8) del análisis de $R_i, i = 1, \dots, 251$. Las diferentes Q-Q plots indican normalidad en los datos, excepto la Figura 6 en la cual se observa el efecto de un punto de cambio.

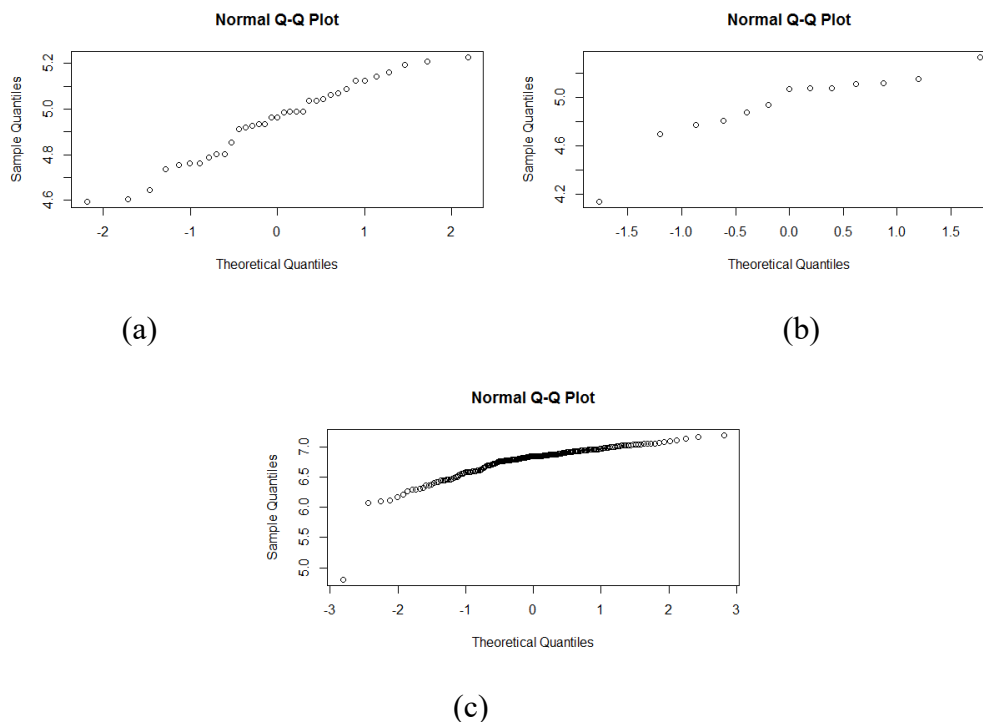


Fig. 8. Q-Q plot de R_i con (a) $i = 1, \dots, 35$, (b) $i = 36, \dots, 48$ y (c) $i = 49, \dots, 251$.

CONCLUSIONES

Los modelos con puntos de cambio son muy útiles en el estudio de cambios en una sucesión de variables aleatorias, esto es importante ya que nos permite tener un mejor ajuste además de que proporcionan información que puede ayudar a entender mejor los fenómenos y tomar mejores decisiones.

Se estudio la cantidad de accidentes de tránsito terrestre en Puebla mediante el modelo normal univariado con puntos de cambio. Las diferentes metodologías indicaron que en los periodos Diciembre del 2000-Enero de 2001 y en Agosto-Septiembre de 2002 se tuvo un aumento significativo en la cantidad de accidentes de tránsito terrestre.

REFERENCIAS

- Chen, J.; Gupta, A. K. (2012). *Parametric Statistical Change Point Analysis*. 2a ed., Birkhäuser.
- Hawkins, D. M. (1977). Testing a sequence of observations for a shift in location. *Journal of the American Statistical Association*, 72, 180-186.
- INEGI (2016). *Síntesis metodológica de la estadística de accidentes de tránsito terrestre en zonas urbanas y suburbanas 2016*. Instituto Nacional de Estadística y Geografía, México.
- Horváth, L. (1993). The maximum likelihood method for testing changes in the parameters of normal observations. *Annals of Statistics*, 21, 671-680.
- Vostrikova, L. J. (1981). Detecting “disorder” in multidimensional random processes. *Soviet Mathematics Doklady*, 24, 55-59.
- Worsley, K. J. (1979). On the likelihood ratio test for a shift in location of normal populations. *Journal of the American Statistical Association*, 74, 365-367.
- Yao, Y. C.; Davis, R. A. (1986). The asymptotic behavior of the likelihood ratio statistics for testing shift in mean in a sequence of independent normal variates. *Sankhyā*, A48, 339-353.

ⁱ Percance vial que se presenta súbita e inesperadamente, determinado por condiciones y actos irresponsables potencialmente previsibles, atribuidos a factores humanos, vehículos preponderantemente automotores, condiciones climatológicas, señalización y caminos, los cuales ocasionan la pérdida prematura de vidas humanas y/o lesiones, así como secuelas físicas o psicológicas, perjuicios materiales y daños a terceros.