BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA

Facultad de Ciencias Físico Matemáticas

ESTIMACIÓN DE LA EVOLUCIÓN DEMOGRÁFICA DEL ESTADO DE PUEBLA BAJO LA TEORÍA ESTABLE ACOTADA

Tesis presentada para obtener el título de:
Licenciatura en Actuaría

Presenta:
José Alejandro Méndoza De Jesús

Directores de Tesis:
Dra. Hortensia Josefina Reyes Cervantes
Mtro. Javier González Rosas

Enero 2021

ESTIMACIÓN DE LA EVOLUCIÓN DEMOGRÁFICA DEL ESTADO DE PUEBLA BAJO LA TEORÍA ESTABLE ACOTADA

"El problema de la población no tiene solución técnica; requiere una extensión fundamental en moralidad."
- Garrett Hardin

A mis domadores de quimeras, sanadores de cisnes heridos, lectores de estrellas, anclas de vida, luces del polo, herederos del canto de la Tierra, tormentas del océano, estrellas de la cuidad, dueños del canto de las golondrinas, habitantes del séptimo cielo, pescadores de sueños, polvo del amanecer, alivio de bestias heridas, sombras del atardecer, ejércitos de fuego y agua, ecos del tiempo, soles del vigésimo tercer verano, trenes y cohetes del Edén.

Al Señor y Señora de la casa Mendoza.

Agradecimientos

Iniciaré esto con el pleno conocimiento de que no me alcanzará este espacio para agradecer a todos y cada uno de a quienes les debo cada paso de este sinuoso, truculento y frecuentemente frío camino llamado vida.
Sin embargo, lo intentaré...

A mis padres y hermanos, por no sabernos comportar cuando nadie nos veía. Por las risas hasta la media noche en esa mesa de cinco ocupantes de la que nadie terminaba por escapar. Por pelear en los momentos difíciles, que en la mayoría de las ocasiones, era nuestra manera de demostrar que nos importaba seguir cuidándonos la espalda. Por cada valor sin fecha de caducidad, por robarle espacio al tiempo, por redimir caídas, por perdonar errores. Por sobrevivir.
Por no soltarnos la mano, por nunca decir no, por no planear nada, por no llegar a tiempo a ningún sitio, por no huir del pasado, por no cerrar los ojos, por no escapar de las tormentas, por no tenerle a la obscuridad, por no tomar jamás el camino fácil. Por seguir caminando, por seguir de pie.

A mis abuelos, por regalarle a todos la mitad de su alma. Por esperar cada sábado para sentarme en esa mesa que, junto a mi estómago, alimentó mi alma.
Por refugiarme de mis fantasmas, por representar mis promesas, por personificar mis oraciones.
Abuelo, escribí esto el día en que te fuiste. Dejaste una marca muy dentro de todos. No te preocupes, no dejaremos que se borre.

A mis amigos, a todos ustedes.
Por inyectarnos color, por recorrer mil millas para encontrarnos, por ser aviones negros en el cielo, por encontrar nuestras similitudes interceptadas en esta carrera, por recuperarnos de cada suplicio juntos, por dar un paso a la vez, por repartirnos la carga, por sonreírnos a la distancia, por cada error al asumir nuestro papel de adultos, por crecer a nuestro ritmo, por prometer que llegaríamos, por nadar contra corriente, por no fingir, por no retroceder. Por ayudarme a crecer.
Sé que algún día habrán recorrido todos los caminos que están destinados a tomar.

A nuestros profesores, por su tenacidad al empujar a cada generación al mundo más allá de la universidad, por compartir sus conocimientos y experiencias, por cuidar que el motivo que nos hizo llegar ahí no se esfumará (claro que, nunca faltó quien

lo pusiera en duda), por traducir un mundo lleno de teoremas y definiciones dentro de esas cuatro paredes, las cuales fueron testigos de lo que sucedía ahí adentro.
Gracias por nutrir la grande vocación y el arte de enseñar.

Agradecimientos especiales a los profesores: M.C. Brenda Zavala, Mtra. Rosalba Mercado y al Dr. Bulmaro Juarez. Por haber visto este proyecto cuando solo era un borrador y nutrirlo con sus valiosas observaciones. Por atender en cada versión de este proyecto y por haber cumplido mis expectativas con sus comentarios y aportes. Espero haber cumplido las suyas.

Al Mtro. Javier Rosas González, por infundir esfuerzo. Por conducir este proyecto y hacerlo crecer con cada apoyo enviado y cada observación recibida. Por haber sido el mecanismo de aprendizaje en este trayecto y coindicir con nuestro tiempo al haber aceptado colaborar. Gracias.

A la Dra. Hortensia Reyes, le debo mucho, por no decir todo.
Por acoger esta idea incluso antes que yo, por no dejar que lo tirara todo, por convertir su cubículo en el refugio perfecto donde los problemas no parecían tan importantes. Sé que muchos antes que yo lo encontraron y algunos más después de mí lo encontrarán.
Gracias por plantar la inclinación a nuestra ya aclamada estadística, por adentrarme en el mundo de las estimaciones y el misterio de adelantarnos al tiempo con los pronósticos, por todas las lecciones que pudo darle a ese foráneo que, por extrañas razones, terminó en ese salón con usted al frente.
Muchas gracias.

Y por último, a todos los nadie que tuvieron que convertirse en alguien por alguien.

Índice general

Índice de figuras

Índice de cuadros

Introducción

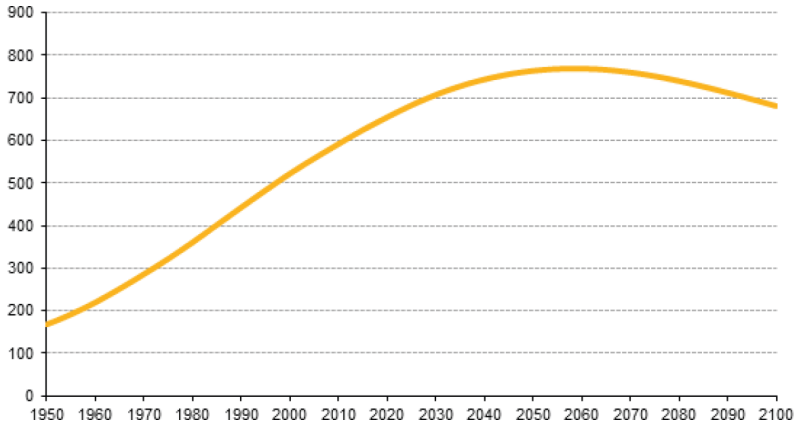
Antecedentes

Desde el inicio de la revolución industrial, el fenómeno demográfico se ha convertido en uno de los mecanismos más influyentes dentro del panorama futuro de cualquier sociedad. Consecuente a guerras, conflictos bélicos y pandemias, la raza humana ha persistido tanto en su permanencia que su reproducción se ha vuelto un factor clave dentro de la historia moderna. Desde años anteriores a los tiempos del demógrafo Thomas Malthus (1766-1834) (Ambientum, 2018), se pronosticó que en algún punto de la historia, las naciones sucumbirían en materia ambiental, energética y sustentable debido al crecimiento masivo de la población. Dicho fenómeno requeriría un mayor desgaste en los aspectos ambientales, económicos, salubres y sociales de cada administración gubernamental, ya que, en caso de no prever la evolución del fenómeno esto desencadenaría una serie de eventos que culminarían con la sobreexplotación ambiental mundial.

Varias organizaciones lideradas por las Naciones Unidas se han unido en la constante evaluación del crecimiento y evolución de la demografía mundial a raíz de la gran magnitud en su importancia. El 11 de julio de 1987 la población mundial rebasó los cinco mil millones de habitantes, dos años después, el Consejo de Administración del Programa de las Naciones Unidas para el Desarrollo estableció esa fecha cómo el Día Internacional de la Población, acontecimiento que enfatiza la importancia del estudio, medición, control y predicción del fenómeno poblacional, así como de todas las variables relacionadas con ella. (Enriquez, 2019). Existen una gran variedad de técnicas y métodos para proyectar poblaciones determinadas en un lugar, entre estos se encuentran modelos de método de componentes que proyectan la población por edad y sexo usando una edad y una estructura de género iniciales y proyecciones de fertilidad, migración y mortalidad. Algunas organizaciones como el Censo Bureau, el Banco Mundial y el Instituto Internacional de Sistemas de Análisis Aplicados, utilizan sus propias metodologías para determinar la tendencia y evolución poblacional (ONU, 2017). Las Naciones Unidas han publicado proyecciones poblacionales usando sus propios métodos basados en tasas futuras de fertilidad, probabilidades de supervivencia y niveles de migración. Sin embargo, no están formalmente sustentados en términos de incertidumbre en términos probabilísticos. Fue hasta el año 2014, cuando la ONU emitió proyecciones probabilísticas por primera vez (Almeka y col., 2015).

El Centro Latinoamericano y Caribeño de Demografía (CELADE), organismo de la división de Población de la Comisión Económica Para América y el Caribe (CEPAL), asumió en 2019 la conceptualización del avance de los distintos factores que contrastan el crecimiento poblacional del área de los países latinoamericanos y el conjunto de islas que conforman el Caribe. Se muestra una desaceleración en el crecimiento del fenómeno población en el área de América Latina y el Caribe (ALC) a partir del año 1990. Durante el quinquenio del año 1950 a 1955 la población en ALC mostró un crecimiento promedio anual de 4.8 millones de personas, mismo que se vio aumentado a 8.2 millones para el quinquenio 1985-1990. A partir de entonces, la población de la región comenzó a sufrir una caída en su tasa de crecimiento. Actualmente, el crecimiento medio anual de población es de 6 millones de habitantes. La figura 1 ilustra la caída de la población en las proyecciones demográficas realizadas por la institución exponiendo un decremento demográfico en el periodo 1950-2100 en millones de personas (CEDALE, 2019).

Figura 1: *Población total proyectada para América Latina y el Caribe.*



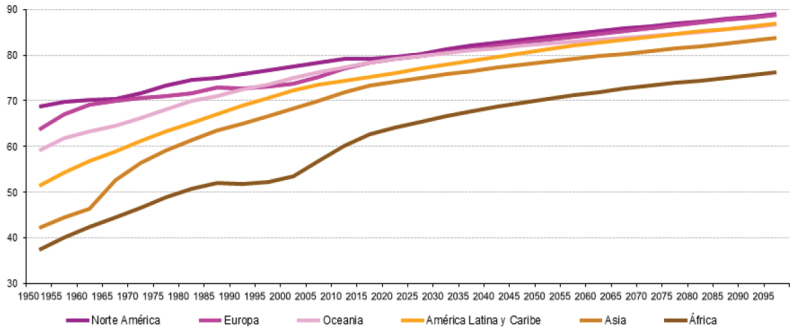
Fuente: CELADE - División de Población de la CEPAL, departamento de Asuntos Económicos y Sociales. Revisión 2019, Naciones Unidas.

Un caso particular se presenta en Cuba puesto que su población ha comenzado a disminuir en el quinquenio actual. Por otro lado, se espera que países como Guatemala y Panamá presenten los últimos lugares de la región en presentar tasas de crecimiento decrecientes, pues se proyecta un crecimiento negativo hasta el periodo 2090-2095. Respecto a la población total, comenzará a disminuir a partir del año 2059. La población máxima de la región en ALC, se alcanzará en el 2058 con un tamaño de 767.5 millones de personas.

La disminución poblacional a largo plazo y la modificación de variables como el nivel de reemplazo y el bono demográfico pueden ser causados por dos principales columnas del estudio demográfico: el aumento en la esperanza de vida y la disminución en la tasa global de fecundidad.(CEDALE, 2019)

La esperanza de vida al nacer en el área seguirá un comportamiento creciente, con un aumento de 0.8 años adicionales al tiempo de vida completo por año durante la última década, América Latina se restringe a ser la penúltima región con el menor incremento de esta variable, solo por arriba del área de Norteamérica que durante los últimos dos quinquenios ha presentado aumentos mínimos en escala global, sin embargo, se compensa con sus elevados niveles de esperanza de vida con un promedio de 79.2 años, la región de Oceanía la sigue de cerca con 78.4 años y Europa con 78.3. Con un valor de 75.2 años completos de vida se encuentra la región de ALC, 73.3 para Asia y debajo de este se encuentra África con 62.7 años.(CEDALE, 2019). La figura 2 ilustra la evolución del tiempo de vida completo por año para cada región de las Naciones Unidas. Se puede apreciar una convergencia latente de esta variable respecto a la línea de tiempo, sin embargo, aún existe una diferencia evidente entre las regiones del primer mundo y las menos favorecidas, esto es debido principalmente a los altos niveles de mortalidad que existe en los lugares más marginado donde los servicios de salud y estabilidad económica aún presentan retrasos.

Figura 2: *Esperanza de vida estimada para las regiones declaradas por ONU.*



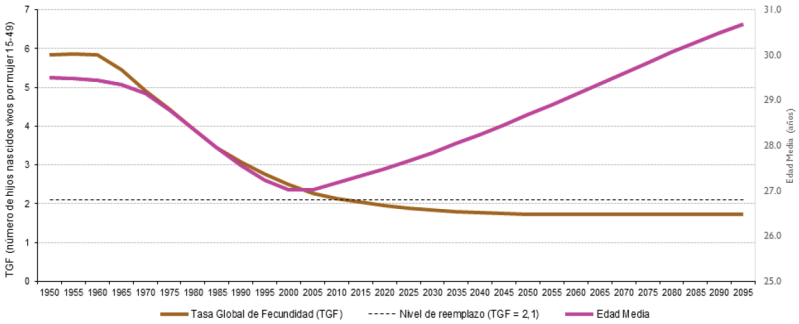
Fuente: CELADE - División de Población de la CEPAL, departamento de Asuntos Económicos y Sociales. Revisión 2019, Naciones Unidas.

La tasa global de fecundidad (TGF) en la región latina y caribeña demuestra un comportamiento descendente. Tan solo en el quinquenio anterior la tasa mostraba un valor promedio de 2.04 hijos por cada mujer en edad reproductiva del área, valor muy distante de los 6 hijos promedio que se observaban hace 70 años. La TGF actual está por debajo del nivel de reemplazo poblacional considerado con el valor de 2.1 hijos por mujer.

Un factor importante para este fenómeno es la edad media de fecundidad. A pesar de sufrir una disminución entre el periodo 1950-2000, ha reportado un crecimiento durante unos años atrás. Hoy en día posee un valor de 27.3 años, lo que implica que una mayor parte de las mujeres de la región están comenzado su periodo reproductivo a edades más avanzadas provocando no solo una disminución de la TGF sino además un periodo de desaceleración para la curva poblacional de la región. El comportamiento de la TGF, así como la edad media de fecundidad, se

observa en la figura 3 donde se comprueba que variables influyentes como la educación sexual en grupos como los adolescentes, el avance tecnológico y el aumento en la escolaridad de la población contribuyen de manera importante al descenso de la TGF.

Figura 3: *Tasa global de fecundidad y edad media de fecundidad para el periodo 1950 - 2100.*



Fuente: CELADE - División de Población de la CEPAL, departamento de Asuntos Económicos y Sociales. Revisión 2019, Naciones Unidas.

El embarazo adolescente (15-19 años) ha mantenido una evolución decreciente. Actualmente 63 de cada mil nacidos vivos pertenecen a mujeres en este rango de edad, a diferencia de los 68.1 que se demostraban en años anteriores. A pesar de su disminución, el área de América Latina y el Caribe, se encuentra en segundo puesto con mayor reincidencia en embarazos adolescentes, solo por debajo del área de África con una tasa de fecundidad de 95 por cada mil nacidos vivos. Regiones como Norteamérica y Europa conservan los últimos lugares con tasas de 18.9 y 12.7 nacimientos por cada mil nacidos vivos, respectivamente. Se presenta de manera puntual las tasas de fecundidad observadas en países con las tasas más elevadas en el área de ALC (Tabla 1), además de presentar su posición a nivel mundial respecto a sus tasas de fecundidad. Podemos apreciar que República Dominicana se encuentra entre los primeros treinta países con mayor recaudación de embarazos adolescentes.

Respecto al panorama nacional, en el año 2018 la encuesta nacional de la dinámica demográfica a cargo de Instituto Nacional de Estadística y Geografía (INEGI), reflejo una población nacional de 125 millones de personas, siendo mujeres el 51.1 % de la población y el 48.9 % restante constando de hombres (Hernández y col., 2013). En un lapso de 225 años (1790-2015), factores sociales, políticos y económicos han influido en el crecimiento demográfico nacional creando tres particiones en la evolución del fenómeno poblacional en México. La primera parte consta del periodo de 1790 a 1930 cuando la población tuvo un crecimiento relativamente lento con tendencia muy cercana a la lineal. El segundo escenario prevalece de 1930 a finales del siglo XX donde el crecimiento se comporta de manera exponencial. A partir del año 2000 el crecimiento comienza a desacelerarse, presentando una disminución en

Tabla 1: *Países con mayores tasas de fecundidad adolescente para el área de ALC. (Hijos nacidos vivos por cada mil embarazos entre las edades 15-19).*

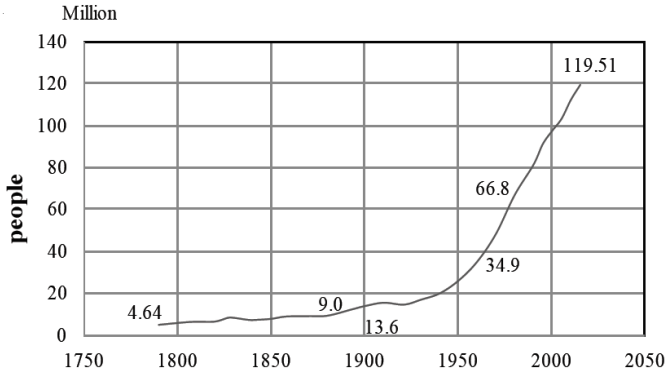
PAÍSES ALC	TASA DE FECUNDIDAD	RANKING MUNDIAL
República Dominicana	94.3	26
Venezuela	85.3	31
Panamá	81.8	34
Ecuador	79.3	35
Guatemala	70.9	46
Paraguay	70.5	47
El Salvador	69.5	48
Colombia	66.7	53
Bolivia	64.9	58
Argentina	62.8	61
México	60.4	65
Brasil	59.1	67
Uruguay	58.7	68
Guyana Francesa	58.4	69
Perú	56.9	71

Fuente: CELADE - División de Población de la CEPAL, departamento de Asuntos Económicos y Sociales. Revisión 2019, Naciones Unidas.

su tasa de crecimiento. Como se observa en la figura 4, en el año 1750 al país le tomó poco más de un siglo (1750-1880) para duplicar su población. Para 1900 solo le tomó 80 años para casi quintuplicar su demografía. Sin embargo, se ha encontrado una disminución en la rapidez del crecimiento del fenómeno pues al considerar el periodo 1980-2015 se observa que la población no ha duplicado su tamaño.

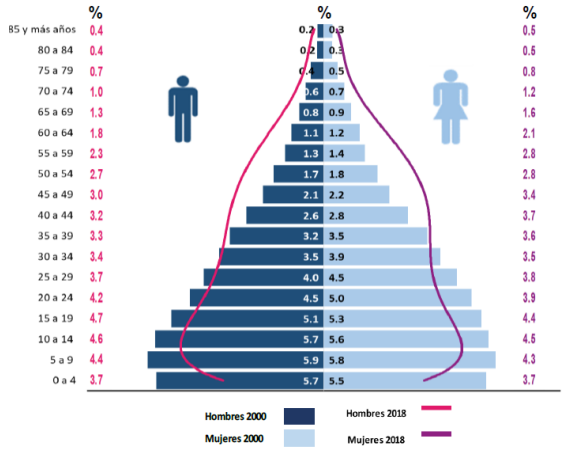
La edad mediana nacional pasó de 22 años a 29 años en el periodo 2010-2018, además, la TGF ha descendido en los últimos años. Se registró un promedio de hijos de 2.7 en el año 2009, mismo que se redujo a 2.4 en la actualidad. En cuanto a la pirámide poblacional se observan cambios importantes en los últimos años. La base de la población mexicana sigue reduciéndose, lo que representa un tamaño menor en la población de grupos infantiles menores de 15 años, la participación de los grupos de edades de 15 a 29 años continuará en descenso mientras que los grupos de 30 años en adelante comenzarán a aumentar en los próximos años. El grupo de 60 años y más sigue mostrando un aumento sustancial pues en el periodo 2000-2018 ha pasado de 7.3 % a 12.3 % de la población total nacional. Se muestra una comparación gráfica (figura 5) entre los resultados obtenidos de la encuesta del año 2000 y la emitida en 2018 por INEGI. Se puede observar una diferencia considerable en la base de la pirámide poblacional a cargo de los grupos de edad de 0-15 años, además de sufrir un aumento en los grupos mayores a los 60 años tanto en hombres como en mujeres, demostrando un latente envejecimiento poblacional (INEGI, 2019).

Figura 4: *Población en México para el periodo 1790-2015.*



Fuente: INEGI. Encuestas Nacionales de la Dinámica Demográfica 1790-2015.

Figura 5: *Comparación de la estructura de la población en México por grupo quinquenal de edad y sexo entre los años 2000 y 2018.*



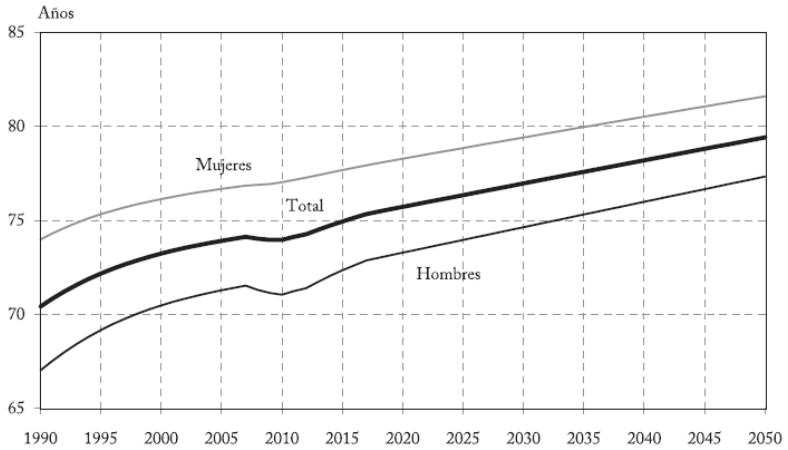
Fuente: INEGI. Encuesta Nacional de la Dinámica Demográfica 2018.

En el 2013, la población joven, los grupos de edad de 12-29 años, representaba aproximadamente un tercio de la población total (32.1 %), mientras que la población de 60 años en adelante, constituía el 9.5 por cierto. Esto implicaba que por cada 3.4 jóvenes existía una persona en edad de vejez. Se proyecta que en 2030 esta relación se torne en 2 por cada adulto mayor de 60 años y será en 2050 cuando esta relación sea uno a uno y el tamaño de sus grupos sean similares en porcentaje poblacional (Hernández y col., 2013).

A finales del año 2000, la esperanza de vida nacional constaba de un valor de 73.2 años. Esta variable demográfica alcanzó los 74.1 en los siguientes años siete años debido a los avances tecnológicos y científicos en materia de salud y las condiciones mejoradas en economía y salubridad del país. Actualmente, la población mexicana registra una esperanza de vida de 74.5 años. La figura 6 compara la diferencia entre la esperanza de vida de hombres y mujeres mexicanos, así como la esperanza vida

media nacional a partir de los registros que datan de 1990 incluyendo las proyecciones emitidas hasta el año 2050.

Figura 6: *Esperanza de vida total mexicana, 1990-2050.*



Fuente: CONAPO. Proyecciones de población 2010-2050.

Un factor relevante del país se centra en la TGF mexicana que se ha modificado de manera notable en los últimos años. La inserción laboral de las mujeres, la educación sobre sexualidad, así como la modificación en las preferencias reproductivas de la población son factores que actualizan el valor del promedio de hijos esperado por las mujeres mexicanas. La Encuesta Nacional de la Dinámica Demográfica (ENADID, 2018), demostró una reducción de la fecundidad de las mujeres. Tan solo en el lapso de 1960 a 2009 la TGF pasó de 7 a 2.7 hijos, en el 2014 resultó en 2.6, para el 2018 esta se redujo a 2.4 hijos en promedio para cada mujer en edad fértil.

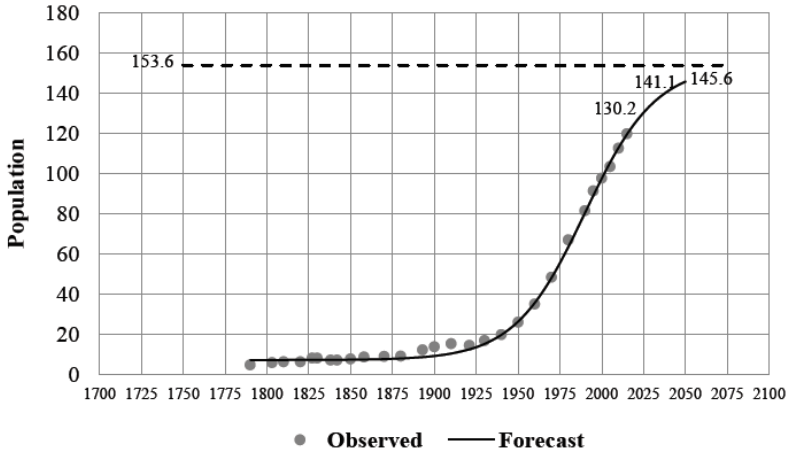
Una variable que puede explicar la reducción de la TFG es el uso de métodos anticonceptivos. Según las ENADID, el porcentaje de mujeres en edad fértil usuarias de métodos anticonceptivos pasó de 72.3 % a 73.1 % en el lapso de 2014 a 2018. Una mayor autonomía y un aumento en su poder de decisión y posición social han ampliado las posibilidades del género femenino en el país (INEGI, 2019).

Bajo la teoría de modelos logísticos, en conjunto con los resultados de la Teoría Estable Acotada (González-Rosas & Zárate-Gutiérrez, 2018),existen cotas poblacionales que determinan un límite para el crecimiento de las poblaciones. En otras palabras, se prueba la existencia de un mínimo y máximo para la demografía. En el caso de México, se demostró la existencia de un mínimo y máximo para la población nacional con valores resultantes en 7.1 y 153.6 millones respectivamente. Esta última cifra demuestra la existencia de un tope para la población nacional que puede ser ocasionado por un sobrecargo en el entorno social, económico y ambiental. Problemas como el deterioro ambiental, altas tasas de desempleo, reducción en las tasas de fertilidad, reducción de espacio geográfico debido al aumento de los mares, el desabasto de alimentos y agua potable en tiempos posteriores son acontecimientos

de delimitan el crecimiento demográfico del país.

La figura 7 presenta de forma gráfica el comportamiento de la población nacional considerando la existencia de restricciones en su entorno. La constante de 153.6 acota a la población por arriba, a la cual de manera asintótica se estará acercando en los próximos años.

Figura 7: *Proyección de crecimiento demográfico en México bajo la Teoría Estable Acotada, 1790-2050.*



Fuente: GLOBAL JOURNALS INC. The stable bounded theory an alternative to projecting populations, 2018 (González-Rosas & Zárate-Gutiérrez, 2018).

Planteamiento y justificación

Ante una emergencia poblacional de escala mundial, los organismos gubernamentales deberán crear e implementar estrategias económicas, sociales y ecológicas para demandar los servicios en términos de seguridad, empleo, educación y salud para las próximas generaciones. El crecimiento demográfico, la esperanza de vida y el índice de natalidad son columnas sobre las que se sustentan las bases de la ciencia demográfica, por lo tanto, la vigilancia constante de estas variables es de suma importancia para la vida económica futura de la sociedad. Programas como el Plan Nacional de Planificación Familiar del Sector Salud son medidas necesarias para disminuir la tasa de crecimiento de la población y de esta manera evitar la saturación de los servicios médicos y los sistemas de ahorro para el retiro de las personas mayores. En busca de nuevas estrategias, el estudio y análisis demográfico se han convertido en herramientas en pro de una mejor toma de decisiones por parte de las autoridades y sociedades que investigan la creación de mejores escenarios para la población.

En Puebla, dada la situación demográfica actual, el comportamiento creciente de

la esperanza de vida y el no lejano envejecimiento de la población, el impacto de las investigaciones demográficas será cada vez más mayor y el requirimiento de conocer e interpretar la evolución poblacional será imprescindible para la toma de decisiones, por lo que nuevos modelos poblacionales bajo influencia de variables de restricciones en el entorno (pérdida de la ecología, aumento del desempleo, disminución en la calidad del aire, aumento en los impuestos, reducción de espacio geográfico, aumento en casos de enfermedades bajamente tratables, escasez de alimento, etc.) serán cada vez más útiles.

Objetivos generales

El objetivo principal del presente trabajo es demostrar la evolución logística de la población del estado de Puebla, la cual estará acotada por una constante poblacional, misma a la que la población se aproximará en el avance sobre el horizonte de tiempo. En base a lo anterior,se espera resolver los siguientes puntos respecto a la población estudiada:

- Demostrar la presencia de un patrón logístico sobre el fenómeno demográfico estatal.
- Probar la existencia de un mínimo y máximo para la población del estado de Puebla.
- Cuantificar el valor numérico del mínimo y máximo poblacional.
- Obtener un modelo matemático para el crecimiento demográfico del estado.
- Realizar los pronósticos para el crecimiento poblacional estatal para las 3 próximas décadas mediante la Teoría Estable Acotada.

Descripción del contenido

A continuación se relata de manera descriptiva el contenido de cada uno de los capítulos del presente trabajo.

En el capítulo uno se abordan los conceptos básicos para definir la metodología de modelos de regresión, sus bases y sus implicaciones. Incluyen anotaciones sobre la diferenciabilidad de funciones y diversos métodos de integración, además de tocar puntos sobre el teorema fundamental del cálculo, la diversidad de las ecuaciones diferenciales y su influencia sobre los modelos de crecimiento logístico.

Debido a la naturaleza del problema, es conveniente recurrir a la estimación de modelos y sus parámetros por lo que se consideran metodologías que permiten su resolución.Dentro de las estas metodologías se consideran principalmente dos de ellas:

la estimación por mínimos cuadrados ordinarios y la estimación por máxima verosimilitud. Por principio de parsimonia, se elige la primera metodología por encima de la segunda debido a que para usar máxima verosimilitud se necesita conocer la distribución de las variables a estudiar, condición que es desconocida para el estudio. La teoría de los modelos de regresiones lineales y múltiples a partir de sus respectivos supuestos se presenta en el capítulo dos y se conceptualiza la metodología de mínimos cuadrados ordinarios, los cuales desempeñan un papel importante dentro del presente estudio. Más tarde, se introducen conceptos de estimación por intervalos para los distintos estimadores que forman parte de los modelos de regresión.

El capítulo tres aborda la aplicación de la teoría estable acotada, así como la teoría de modelos de crecimiento logístico específicamente sobre la demografía del estado de Puebla, se utiliza la teoría para determinar un modelo matemático que simule la demografía estudiada y posteriormente se pronostican las medias poblacionales para la siguiente década en el estado.

El análisis de los resultados y las conclusiones obtenidas en el capítulo tres se concentran en el último capítulo. De manera descriptiva y gráfica se muestran una comparación entre los pronósticos obtenidos y los datos que componen la muestra demográfica, se enfatiza en las causas y consecuencias del fenómeno y la progresión en las decisiones que se deben considerar respecto al crecimiento poblacional.

Finalmente, el apartado de los apéndices se enfoca en las distintas pruebas y conceptos estadísticos para la valoración y evaluación de los diferentes supuestos que deben respetarse sobre los modelos de regresiones múltiples y lineales. Conceptos como la prueba T, Análisis de Varianza, el coeficiente de determinación, el valor P y las pruebas Shapiro-Wilks, White, y Durbin-Watson validan el presente trabajo. Además, se incluye la solución de la ecuación diferencial para modelos demográficos con cotas poblacionales y por último se anexa el código utilizado en lenguaje del software estadístico R donde se realizaron las gráficas y cálculos presentados en esta obra.

Capítulo 1

Preliminares

La estadística nos brinda herramientas que se encuentran en una amplia gama de áreas dentro de las matemáticas aplicadas para la modelación de situaciones observables. Una de ellas es el estudio y la estimación por medio de las regresiones lineales bajo la metodología de mínimos cuadrados ordinarios (MCO), la cual sustenta sus bases en las leyes de la probabilidad y la estadística. La modelación matemática reside en elementos indispensables como la formulación e interpretación del cálculo diferencial e integral, ecuaciones diferenciales, además del uso de herramientas en probabilidad y estadística, trabajando en conjunto con el uso de la programación (lenguajes matemáticos y paquetes estadísticos). Estos y más temas implícitos son requeridos para el desarrollo de instrumentos sólidos de predicción para la obtención de resultados de los fenómenos a estudiar.

Para el presente trabajo se inicia con la introducción de conceptos básicos que amplían la comprensión y la construcción matemática de la línea de investigación que se presenta, los cuales se resumen en funciones integrables y sus propiedades, el concepto de integración y el método de integración por fracciones parciales, introducción al tema de ecuaciones diferenciales y el método sistematizado de variables separables así como el análisis de los modelos de población logísticos.

1.1. La integral de Riemann

Definición 2.1: (Perez-J, 2008) *Sea* $f : [a, b] \to R$ *una función acotada y sea* $P = \{t_0, t_1, \ldots t_n\}$ *una partición del intervalo [a,b]. Para cada* $k \in \{1, 2, \ldots, n\}$ *llamaremos* I_k *al intervalo* $[x_{k-1}, x_k]$ *y denotemos los siguientes conceptos:*

$$M_K(f, P) = \mathrm{Sup} f(I_k).$$

$$m_k(f, P) = \mathrm{Inf} f(I_k).$$

Llamaremos ***suma superior de Riemann de la función f respecto a la partición P*** *al número real:*

$$S(f,P) := \sum_{k=1}^{n} M_k(f,P)\,(x_k - x_{k-1})\,.$$

Análogamente, llamaremos ***suma inferior de Riemann de la función f respecto a la partición P*** *al número real:*

$$I(f,P) := \sum_{k=1}^{n} m_k(f,P)\,(x_k - x_{k-1})\,.$$

Cumpliéndose siempre que $I(f,P) \leq S(f,P)$.
Llamaremos al ínfimo del conjunto de sumas superiores ***integral superior de Riemman de f*** *en el intervalo [a,b] denotándolo por:*

$$\overline{\int_a^b} f(x)\,\mathrm{d}x.$$

Por otro lado, llamaremos al supremo del conjunto de sumas inferiores ***integral inferior de Riemann de f*** *en el intervalo [a,b] denotándolo por:*

$$\underline{\int_a^b} f(x)\,\mathrm{d}x.$$

Diremos que ***f es integrable Riemann en el intervalo [a,b]*** *si* $Inf(S) = Sup(I)$*, es decir, si se cumple que:*

$$\overline{\int_a^b} f(x)\,\mathrm{d}x = \underline{\int_a^b} f(x)\,\mathrm{d}x.$$

Este valor será conocido como la ***integral de Riemann de f en [a,b]*** *y se denota por:*

$$\int_a^b f(x)\,\mathrm{d}x.$$

El anterior resultado demuestra que mientras una función f sea acotada en un intervalo dado, entonces f es integrable en el intervalo dado si y solo si existe una sucesión de particiones $\{P_n\}$ verificando que la sucesión $\{S(f,P_n) - I(f,P_n)\}$ converge a cero.

1.1.1. Propiedades de las funciones integrables

Se citan las siguientes propiedades de las funciones que cumplen con la condición de ser integrables (Colegio-de-México, 2012).

Proposición 2.1: Sean $f, g : [a,b] \to R$ dos funciones integrables en $[a,b]$. Entonces:

1. $f + g$ es una nueva función integrable en $[a, b]$ y se verifica que:

$$\int_a^b (f+g)(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x + \int_a^b g(x)\,\mathrm{d}x.$$

2. Para cada $r \in R$, la función f es una nueva función integrable en $[a, b]$, se verifica que:

$$\int_a^b (rf)(x)\,\mathrm{d}x = r\int_a^b f(x)\,\mathrm{d}x.$$

3. Si para cada $x \in [a, b]$, $f(x) \leq g(x)$, se tiene que:

$$\int_a^b f(x)\,\mathrm{d}x \leq \int_a^b g(x)\,\mathrm{d}x.$$

1.1.2. Teorema Fundamental del Cálculo

Para enunciar el teorema fundamental del cálculo se introduce el concepto de integral indefinida.

Definición 2.2: (Colegio-de-México, 2012) *Sea I un intervalo de números reales y una función continua $f : I \to R$. Si $c \in I$, se conoce como* ***integral indefinida de f con origen en c*** *a la función $F : I \to R$, definida, para cada $x \in I$, como:*

$$F(x) = \int_c^x f(t)\,\mathrm{d}t.$$

Teorema 2.1 ***(Teorema fundamental del cálculo).*** *Sea f una función continua en un intervalo I y sea F cualquier integral indefinida de f. Entonces F es derivable en I y para cada $x \in I$,*

$$F'(x) = f(x).$$

1.1.3. Métodos de integración

Definición 2.3: *Sea una función f definida en un intervalo I se dice que admite una* ***función primitiva*** *si existe una función $G : I \to R$ derivable tal que para todo $x \in I$, $G'(x) = f(x)$.*

Un resultado consecuente del teorema del valor medio permite evaluar la integral de una función conocida su primitiva.

Teorema 2.2: *(Regla de Barrow). Sea $f : [a, b] \to R$ una función integrable y suponemos que admite una función primitiva G. Entonces,*

$$\int_a^b f(x)\,\mathrm{d}x = G(b) - G(a).$$

En muchos casos es conveniente transformar la función f en otra función cuya primitiva sea más accesible.

Corolario 2.1: *(Teorema de cambio de variable) Sea $g : [a, b] \to R$ una función de clase $C^1([a, b])$ con $g'(x) \neq 0$. Si f es una función continua en $g[(a, b)]$, entonces la función $f \circ g'$ es una nueva función integrable y*

$$\int_{g(a)}^{g(b)} f(x)\,\mathrm{d}x = \int_a^b f(g(t)) * g'(t)\,\mathrm{d}t.$$

La siguiente técnica es útil cuando se trata de calcular la integral de un producto de funciones, o bien, de una función que puede ser vista como el producto de dos funciones más.

Corolario 2.2: (Colegio-de-México, 2012) *(Teorema de integración por partes). Sea $F, G : [a, b] \to R$ dos funciones de clase* C^1 *$([a, b])$. Entonces:*

$$\int_a^b F(x)G'(x)\,\mathrm{d}x = F(b)G(b) - F(a)G(a) - \int_a^b F'(x)G(x)\,\mathrm{d}x.$$

Corolario 2.3: (Del-Pino, 2019) *(Integración por fracciones parciales) Sea $f : I \to R$ una función racional (cociente de dos polinomios) y sean $P, Q : R \to R$ las correspondientes funciones polinómicas tales que $f(x) = \frac{P(x)}{Q(x)}$ con $Q(x) \neq 0$ para $\forall$ $x \epsilon I$. Entonces $f(x)$ puede ser expresado de la forma:*

$$f(x) = \frac{P(x)}{Q(x)} = T(x) + \frac{R(x)}{Q(x)}.$$

Donde $T(x)$ es el polinomio resultante de la división y $R(x)$ es el resto de la división cumpliéndose siempre que el grado de R es menor que el divisor $Q(x)$. De esta forma toda función racional se puede escribir como la suma de un polinomio y una función racional propia. Por otro lado, toda función racional propia puede descomponerse en suma de fracciones de la forma:

$$\frac{A}{(\alpha x + \beta)^K},$$

y

$$\frac{Bx + C}{(ax^2 + bx + c)^m}.$$

Donde:

- k, m ϵN.

- a, b, c, A, B, C,α,β son constantes.

Entonces el cálculo de la integral de una función racional se reduce al cálculo de integrales de polinomios y a cálculos de integrales de la forma:

$$\int \frac{A}{(\alpha x+\beta)^K}\,\mathrm{d}x,$$

y

$$\int \frac{Bx+C}{(ax^2+bx+c)^m}\,\mathrm{d}x.$$

1.2. Ecuaciones diferenciales de variables separables

Definición 2.4: *Se define a una* ***ecuación diferencial*** *como una función que incluye una variable dependiente y sus derivadas, con respecto a una o más variables independientes. Si la ecuación contiene derivadas respecto a una sola variable independiente entonces se dice que es una ecuación diferencial ordinaria (EDO) (Bravo, 2017), (Zill, 2011).*
El **orden**(n) de una ecuación diferencial es el valor más alto de la derivada que aparece en la ecuación, mientras que el **grado** es el mayor exponente de la derivada de mayor orden.

Definición 2.5: *Se llama solución de la EDO de orden n en un intervalo I a toda función real $f : I \to \mathrm{R}$ que verifica las siguientes propiedades:*

- *f es n-veces derivable en I.*
- *Para cada $x \epsilon I$, se verifica la ecuación cuando se sustituye la incógnita por la función f.*

Definición 2.6 *Un* ***problema de valor inicial*** *es una ecuación diferencial (ED) que está enlazada con condiciones iniciales. Para verificar que una función es solución de la ED se debe confirmar que satisface las condiciones iniciales.*

Antes de tratar de encontrar una solución particular a un problema de valor inicial es necesario verificar su existencia, y de ser así, debe ser única. Para el caso de ecuaciones de primer orden se obtiene lo siguiente.

Teorema 2.3: *Sea $y' = f(x, y)$, $y(x_0) = y_0$, un problema de valores iniciales. Si existe un rectángulo R del plano tal que $(x_0, y_0) \epsilon R^{int}$ y verificando que f, $\frac{\partial f}{\partial y} \epsilon C(R)$, entonces existe un intervalo $I = (x_0 - \delta, x_0 + \delta)$, $\delta \geq 0$ y una única función $y = f(x)$ definida en I tal que la solución de la EDO que verifica la condición adicional $y_0 = f(x_0)$.*

1.2.1. Ecuaciones diferenciales ordinarias lineales de primer orden

Definición 2.7 *Una ecuación diferencial ordinaria se dice que es* ***lineal de orden uno o de primer orden*** *si es de la forma:*

$$y' + a(x)y = b(x).$$

Donde $b, a : I \to R$*, son dos funciones continuas definidas en un intervalo* I *de números reales. Si* $b(x) = 0$*,* $\forall\ x \epsilon\ I$*, se dice que dicha ecuación es* ***homogénea****. Si* $b(x) \neq 0$*, entonces la ecuación es* ***no homogénea.***

- Caso homogéneo: Es claro ver que la ecuación homogénena de primer orden puede escribirse de la forma:

$$\frac{y'(x)}{y(x)} = -a(x),$$

o bien,

$$\log(y(x)) = -A(x).$$

Donde $A(x)$ es una primitiva de la función $a(x)$. Entonces una solución de la ecuación es de la forma:

$$f(x) = Ce^{-A(x)}, C\epsilon\mathrm{R}.$$

Si suponemos la condición adicional $y(x_0) = y_0$, la solución es única al ser determinada por la primitiva.

- Caso no homogéneo:La función de la forma:

$$f(x) = (C + B(x))e^{-A(x)}.$$

Es una solución de la ecuación $y'+a(x)y = b(x)$ donde $C\ \epsilon$ R, $A(x)$ es cualquier primitiva de $a(x)$ y $B(x)$ es cualquier primitiva de $b(x)e^{A(x)}$. Análogamente, si fijamos una condición adicional, la solución quedará determinada de forma única.

1.2.2. Ecuaciones diferenciales ordinarias no lineales de primer orden

Dentro de esta clasificación de EDO se permite la existencia de distintos tipos de ecuaciones. A continuación, se presentan algunos de ellos (Bravo, 2017),(Zill, 2011).

TIPO I: Ecuaciones diferenciales de primer orden homógeneas.

Definición 2.8: *Una EDO de primer orden se dice que es* ***homogénea*** *si es de la forma:*

$$y' = F\left(\frac{y}{x}\right).$$

Con x diferente de cero, y ϵ R , siendo f una función continua en cierto conjunto $D \subseteq R^2$. Una función f es solución si, y sólo si $\frac{f(x)}{x}$ es solución de la ecuación separada $u' = \frac{F(u)-u}{x}$.

TIPO II: Ecuaciones diferenciales de variables separadas.

Definición 2.9: *Una EDO de primer orden se dice que es de variables separables si es de la forma:*

$$y' = G(t, y).$$

Donde

$$G = P(t)Q(y).$$

Además, $P(t)$ y $Q(y)$ son dos funciones continuas en intervalos I y J con $Q(y)\neq 0$ $\forall$ y ϵ J.

Se tiene que f definida en el intervalo I y es solución de la ecuación anterior si satisface la siguiente igualdad:

$$B(f(t)) = A(t).$$

Donde B es una primitiva de la función $\frac{1}{Q(y)}$ y A es una primitiva de $P(t)$. El método de separación de variables es una técnica que reduce el problema de solucionar ciertas ecuaciones diferenciales ordinarias de primer orden a evaluar dos integrales. La separación de variables funciona cuando se puede escribir la ecuación diferencial en la forma:

$$\frac{\mathrm{d}y}{\mathrm{d}t} = P(t)Q(y).$$

Definición 2.10 *Método por separación de variables(Bravo, 2017).*
Procedimiento sistematizado:

1. Comprobar que la ecuación diferencial es EDO de primer orden. Supongamos que la función a resolver es $y = y(t)$.

2. Expresar $\frac{\mathrm{d}y}{\mathrm{d}t}$ como función de t e y solamente:

$$\frac{\mathrm{d}y}{\mathrm{d}t} = h(t, y).$$

3. De ser posible, reescribir $h(t, y)$ como producto de dos funciones, una dependiente de la variable t y la otra en función de y:

$$h(t, y) = g(t)f(y).$$

4. Separar las variables t e y. Usar operaciones elemental para dejar en la parte derecha la variable t y en la izquierda la variable y, tal que resulte:

$$\frac{\mathrm{d}y}{f(y)} = g(t)\mathrm{d}t.$$

Se debe verificar si $f(y) \neq 0$, ya que esto inválida los cálculos. En caso de obtener dicho resultado, podría incluirse en la lista de soluciones.

5. Integrar ambos lados para obtener una ecuación de la forma:

$$F(y) = G(t) + c.$$

6. Si es posible, resolver y en términos de t.

7. Verifique si es posible incluir soluciones de la división por cero (paso 4). De ser así, incluirla junto con las soluciones obtenidas en este paso, juntas conforman la solución general.

8. Verificar los resultados, es decir, que la solución general cumpla con la ecuación diferencial original

1.3. Modelos poblacionales logísticos

Los modelos logísticos son usados frecuentemente para estudiar fenómenos poblacionales en especies que demuestran un crecimiento exponencial a partir de ciertas restricciones en su entorno (Georgia-Institute-of-Technology, 2014).

Un modelo más preciso postula que la tasa relativa de crecimiento decrece cuando la población P se aproxima a la capacidad K del entorno para soportar a la población. La ecuación correspondiente bajo el anterior supuesto es llamada **ecuación diferencial logística** y se plantea de la siguiente manera:

$$\frac{\mathrm{d}P}{\mathrm{d}t} = rP\left(1 - \frac{P}{K}\right).$$

Donde:

- $P > 0$, define al tamaño de la población.
- $K > 0$, cuantifica la capacidad del entorno que soporta a la población.
- $r \in (0, 1)$, representa la tasa relativa de crecimiento.

1.3.1. Solución de la ecuación logística

Definición 2.11: La ecuación logística se presenta a través de una ecuación diferencial, por lo tanto se debe encontrar una solución que la satisfaga. Al ser una ecuación diferencial ordinaria de primer orden, se resuelve a través del método de variables separables (definición 2.10). Se puede observar que:

$$\frac{\mathrm{d}P}{\mathrm{d}t} = rP\left(1 - \frac{P}{K}\right).$$

Utilizando el procedimiento sistemático del método de variables separables:

$$\frac{\mathrm{d}P}{\mathrm{d}t} = rP\left(1 - \frac{P}{K}\right).$$

$$\frac{\mathrm{d}P}{P\left(1 - \frac{P}{K}\right)} = r\,\mathrm{d}t.$$

Respetando la consistencia de la igualdad:

$$\int \frac{1}{P\left(1 - \frac{P}{K}\right)}\,\mathrm{d}P = \int r\,\mathrm{d}t.$$

Usando operaciones elementales llegamos a:

$$\frac{1}{P\left(1 - \frac{P}{K}\right)} = \frac{1}{P\left(\frac{K-P}{K}\right)} * \frac{K}{K} = \frac{K}{P(K-P)} = \frac{1}{P} + \frac{1}{K-P}.$$

Sustituyendo,

$$\int \left(\frac{1}{P} + \frac{1}{K-P}\right)\mathrm{d}P = \int \frac{1}{P}\,\mathrm{d}P + \int \frac{1}{K-P}\,\mathrm{d}P = \int r\,\mathrm{d}t.$$

Ingresando funciones primitivas:

$$\ln|P| - \ln|K-P| = tr + C.$$

Por reglas de la función logaritmo:

$$\ln\left|\frac{P}{K-P}\right| = tr + C.$$

Buscamos encontrar una solución para la variable de población, por lo que se llega la función:

$$P = \frac{K}{1 + Ce^{-rt}}.$$

Donde,

$$C = \frac{K - P_0}{P_0}.$$

Con esto, se demuestra la existencia de una solución para el problema de ecuaciones diferenciales en contexto de crecimiento demográfico, además de presentar las condiciones en caso de tener un problema de valor inicial (Georgia-Institute-of-Technology, 2014).

En materia biológica, la mayoría de las poblaciones en el mundo excede su capacidad de carga, desencadenando eventos como el crecimiento de su la tasa de mortalidad y la disminución de su tasa de natalidad, disminuyendo el tamaño poblacional hasta su capacidad de carga o incluso por debajo de ésta.(OpenStax, 2018) Cualquier especie que se enfrente a una sobrecarga en su población solo tiene dos caminos: la adaptación o la extinción de la misma.
¿La raza humana presenta la capacidad de regular su crecimiento y rendirse ante los mecanismos existentes en la naturaleza?

Capítulo 2

Análisis de Regresión Lineal

El objetivo básico del estudio de fenómenos aleatorios a través del método de regresión lineal consiste en establecer y cuantificar la asociación lineal entre una variable dependiente aleatoria y una o más variables independientes no aleatorias que repercuten e integran el fenómeno a estudiar (Canavos, 1988).

A continuación se presentan algunos conceptos de Regresión Lineal y se expande la metodología de mínimos cuadrados ordinarios (MCO).

2.1. El modelo de regresión lineal

En su forma más general y abstracta, el modelo de regresión lineal puede representarse como:

$$Y = f(x_1, x_2, x_3, \ldots, x_k).$$

Donde Y es la variable cuyo comportamiento se pretender explicar a través de las distintas variables $(x_1, x_2, x_3, \ldots, x_k)$ que se suponen potencialmente relevantes como factores explicativos de la primera. El vector denota una lista de parámetros que recogen la magnitud con que las variaciones en los valores de las variables x_i $i = 1, \ldots, n$ se transmiten a variaciones en la variable Y (Novales, 2010).

Los modelos de relación o modelos de regresión lineales son del tipo:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \cdots + \beta_k x_k. \tag{2.1}$$

En el que resulta evidente que los parámetros transmiten directamente efectos inducidos por los valores de las variables x_i sobre la variable Y, que se pretende explicar.

Para el caso particular donde la variable Y se explica a través de una sola variable conocida x el modelo recibe el nombre de **modelo de regresión simple**.

Un modelo de regresión simple exige una relación lineal entre las dos variables que interactúan por lo que tanto se considera la siguiente recta ajustada:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x. \tag{2.2}$$

Donde x es llamada la **variable independiente o predictoria** y y es una variable conocida que recibe el nombre de **variable dependiente o variable respuesta**.

Los parámetros de la recta ajustada se definen como:

- β_1 es la pendiente de la recta. Este concepto es de suma importancia en un modelo de regresión simple ya que establece el nivel de relación entre la variable independiente y la dependiente. Si el parámetro toma valores muy cercanos a 0 indicará una relación débil entre las variables, con valores altos (positivos o negativos) establecerá una relación fuerte entre estas.
- β_0 es el intercepto de la línea u ordenada al origen. Puede tomar interpretaciones diferentes dependiendo el modelo que este siendo considerado. En muchos casos representa el mínimo valor que puede tomar la variable dependiente.

Para observar la relación lineal entre un conjunto de datos se observan las parejas ordenadas $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$, donde asumimos que como función de x_i cada y_i es generado por la función lineal $y = \beta_0 + \beta_1 x$ que será evaluada en x_i y se obtendrá ruido gaussiano (normal).

Es decir, se obtendrá la relación:

$$y = \beta_0 + \beta_1 x + \epsilon_i. \tag{2.3}$$

Donde la nueva variable ϵ_i representa el error aleatorio (residuo) que se obtiene de la muestra. El error se distribuirá bajo una distribución normal con media 0 y varianza constante σ^2:

$$\epsilon_i \sim \mathcal{N}(0, \sigma^2).$$

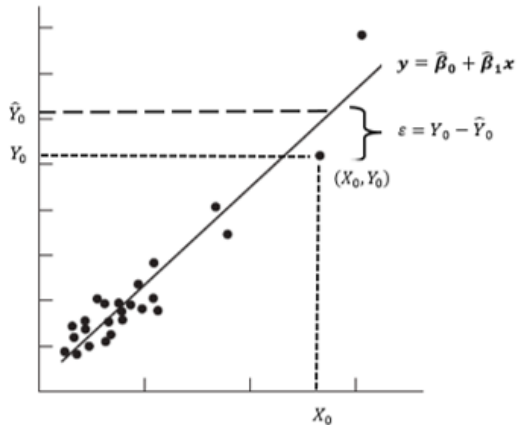
Los valores: β_0,β_1 y σ^2 son los valores poblaciones de los parámetros de la regresión, es decir, son valores desconocidos, por lo que deben ser estimados a través de alguna metodología. Esta estimación debe garantizar la minimización de los n residuos para poder obtener la estimación más acertada posible para los parámetros de la regresión lineal (Novales, 2010).

La figura 2.1 muestra de manera visual la estimación un modelo ajustado sobre un proceso aleatorio considerando parejas ordenadas de datos y los errores aleatorios.

2.2. Supuestos del modelo de Regresión Lineal

Para que un fenómeno pueda ser evaluado por la metodología de regresión lineal debe cumplir con requisitos y reglas que son conocidos como los supuestos de la regresión lineal. Dichos supuestos deben cumplirse para que las estimaciones y evaluaciones sean correctas y no exista evidencia de mal interpretación ni deficiencia teórica (Canavos, 1988), (Carter y col., 2011),(Gujarati & Porter-D., 2010).

Figura 2.1: *Nube de puntos, recta de regresión, valores ajustados y residuos.*



Fuente: Análisis de regresión (Novales, 2010).

Supuesto 1: Los valores de y, para cada valor de x, es

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + e.$$

Supuesto 2: Las variables independientes x_i no son aleatorias y deben tomar al menos dos valores diferentes. En otras palabras, debe existir variabilidad en la muestra.
Supuesto 3: La suma de los residuales, así como su esperanza matemática es nula.

$$\sum_{i=1}^{n} e_i = 0.$$

$$E(e) = 0.$$

Lo dicho anteriormente es equivalente a asumir que:

$$E(y) = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \cdots + \hat{\beta}_k x_k.$$

Supuesto 4: La covarianza muestral entre los regresores y los residuales estimados por MCO es cero. Esto puede ser escrito en términos de los residuales como:

$$\sum_{i=1}^{n} x_i e_i = 0.$$

$$\text{cov}(e_i, x_i) = 0.$$

La media muestral de los residuales obtenidos es cero (supuesto 2), así que, el lado izquierdo de la expresión anterior es proporcional a la covarianza muestral entre x_i y e_i.

Supuesto 5: La varianza de la variable aleatoria del error e es constante para cualquier observación.

$$\text{var}(e) = \sigma^2 = \text{var}(y).$$

Las variables aleatorias y y e tienen la misma varianza porque sólo difieren por una constante. Esto se conoce como *homocedasticidad*.

Supuesto 6: El punto $(\overline{x}, \overline{y})$ esta siempre en la recta estimada por MCO. Es decir, si evaluamos la recta estimada en el punto $\overline{x}$, el resultado será $\overline{y}$.

Supuesto 7: La covarianza muestral entre los valores ajustados $\hat{y}_i$ y los residuales obtenidos es nula. Si reescribimos a cada observación y_i como su valor estimado más su residual:

$$y_i = \hat{y}_i + \hat{e}_i.$$

Por el supuesto 2, la media muestral de los residuales es cero, equivalentemente, la media muestral de los valores estimados $\hat{y}_i$ coincide con el promedio de los y_i. Usando los supuestos 2 y 3 se puede demostrar que la covarianza muestral entre $\hat{y}_i$ y $\hat{e}_i$ es cero. Por lo tanto, los valores ajustados y los residuales obtenidos de la muestra no están correlacionados.

$$\text{cov}(e_i, \hat{y}_i) = 0.$$

Supuesto 8: La covarianza entre cualquier par de errores aleatorios e_i, e_j es nula (ausencia de autocorrelación).

$$\text{cov}(e_i, e_j) = \text{cov}(y_i, y_j) = 0.$$

Este supuesto es importante pues implícitamente sustenta que los residuales e son estadísticamente independientes entre sí.

Supuesto 9: Los residuales e siguen una distribución normal con media 0 y varianza constante.

$$e_i \sim \mathcal{N}(0, \sigma^2).$$

2.3. Estimadores de Mínimos Cuadrados Ordinarios

El estimador de mínimos cuadrados ordinarios utiliza como criterio la minimización de la Suma de los Cuadrados de los errores (SCE), también llamada suma residual.

$$\min_{\beta_0, \beta_1} SCE = \sum_{i=1}^{n} \epsilon_i^2 \tag{2.4}$$

Donde cada residuo asociado a cada observación $i = 1, 2, \ldots, n$ depende los valores de coeficientes escogidos, ya que:

$$\hat{\epsilon}_i = y_i - \left(\hat{\beta}_0 + \hat{\beta}_1 x_i\right) \tag{2.5}$$

Sean $(x_1, y_1), \ldots, (x_n, y_n)$ las parejas ordenadas que constituyen los datos observados, entonces podemos encontrar una línea con la menor distancia entre las observaciones y las estimaciones, que se plantea en resolver el sistema de optimización:

$$\min_{\beta_0,\beta_1} SCE = \sum_{i=1}^{n} \epsilon_i^2 = \min_{\beta_0,\beta_1} \sum_{i=1}^{n} \left(y_i - \left(\hat{\beta}_0 + \hat{\beta}_1 x_i \right) \right)^2 . \tag{2.6}$$

Donde $\min_{\beta_0,\beta_1}$ se interpreta como la minimización de ambos parámetros. Esto es conocido como **modelo de regresión lineal por mínimos cuadrados ordinarios**. Derivando SCE con respecto a ambas variables (β_0, β_1) e igualando dichas derivadas a cero, tenemos:

$$\frac{\partial SCE}{\partial \beta_0} = -2 \sum_{i=1}^{N} \left(y_i - (\beta_0 + \beta_1 x_i) \right). \tag{2.7}$$

$$\frac{\partial SCE}{\partial \beta_1} = -2 \sum_{i=1}^{N} \left(y_i - (\beta_0 + \beta_1 x_i) \right) x_i. \tag{2.8}$$

Con matriz de segundas derivadas:

$$\frac{\partial^2 SCE}{\partial \beta_0 \partial \beta_1} = \tag{2.9}$$

$$\begin{pmatrix} 2n & 2\sum_{i=1}^{n} x_i \\ 2\sum_{i=1}^{n} x_i & 2\sum_{i=1}^{n} x_i^2 \end{pmatrix}.$$

Que tiene por determinante:

$$DET = 4 \left(n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i \right)^2 \right) = n^2 \left(\frac{\sum_{i=1}^{n} x_i^2}{n} - \overline{x}^2 \right) = n^2 \frac{\sum_{i=1}^{n} (x_i - \overline{x}^2)}{n} = n^2 S_x^2.$$

Dado un conjunto de datos y las ecuaciones normales dadas por 2.7 y 2.8, se procede a resolver el sistema:

$$\sum_{i=1}^{n} y_i = n\beta_0 + \beta_1 \sum_{i=1}^{n} x_i. \tag{2.10}$$

$$\sum_{i=1}^{n} y_i x_i = \beta_0 \sum_{i=1}^{n} x_i + \beta_1 \sum_{i=1}^{n} x_i^2. \tag{2.11}$$

Despejando β_0 de 2.10, se tiene que:

$$\hat{\beta}_0 = \frac{\sum_{i=1}^{n} y_i - \beta_1 \sum_{i=1}^{n} x_i}{n} = \overline{y} - \hat{\beta}_1 \overline{x}. \tag{2.12}$$

Sustituyendo 2.12 en 2.11 y despejando β_1, se obtiene:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i x_i - \frac{1}{n}(\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i)}{\sum_{i=1}^{n} x_i^2 - \frac{1}{n}(\sum_{i=1}^{n} x_i)^2} = \frac{\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{n} (x_i - \overline{x})^2} = \frac{S_{xy}}{S_x^2} = \rho_{xy} \frac{S_y}{S_x}. \tag{2.13}$$

Donde ρ_{xy} es llamado el coeficiente de correlación y se define por la relación:

$$\rho_{xy} = \sum_{i=1}^{n} \left(\frac{x_i - \overline{x}}{S_x} \right) \left(\frac{y_i - \overline{y}}{S_y} \right) = \frac{S_{xy}}{S_x S_y}. \tag{2.14}$$

Además, S_{xy},S_x^2,S_y,S_x denotan la covarianza, varianzas y desviaciones estándar muestrales de X y Y. Las estimaciones de MCO para los parámetros son dadas por 2.12 y 2.13 como función de estadísticos muestrales.

2.4. Intervalos de confianza

Para la obtención de pruebas de hipótesis y calcular intervalos de confianza, es necesario conocer con antelación nuestro estadístico de prueba, su desviación estándar, así como su distribución. Todos los estadísticos de prueba presentados se basan en el supuesto de que el fenómeno a estudiar puede expresarse como la ecuación 2.1.

2.4.1. Pendiente

Para la pendiente β_1, el estadístico de prueba está dado por:

$$t_{\beta_1} = \frac{\hat{\beta}_1 - \beta_1}{S_{\beta_1}}. \tag{2.15}$$

El cuál tiene una distribución *T de Student* (Apéndice C) con $n-2$ grados de libertad (Montgomery y col., 2012). El error estándar de la pendiente S_{β_1} se define como:

$$S_{\beta_1} = \frac{\overline{\sigma}}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2}}. \tag{2.16}$$

Y el error cuadrado medio $\overline{\sigma}^2$ es:

$$\overline{\sigma}^2 = \frac{\sum_{i=1}^{n}(\hat{y}_i - \overline{y})^2}{n-2}. \tag{2.17}$$

El intervalo de confianza para el estimador de β_1 se establece como:

$$IC = \left(\hat{\beta}_1 - t_{(\frac{\alpha}{2}, n-2)} S_{\beta_1}, \hat{\beta}_1 + t_{(\frac{\alpha}{2}, n-2)} S_{\beta_1} \right). \tag{2.18}$$

Donde n son los grados de libertad y α es el nivel de significancia de la prueba.

2.4.2. Intercepto

Para el intercepto β_0, el estadístico de prueba está dado por:

$$t_{\beta_0} = \frac{\hat{\beta}_0 - \beta_0}{S_{\beta_0}}. \tag{2.19}$$

El cuál también posee una distribución *T de Student* con $n-2$ grados de libertad. El error estándar de la pendiente S_{β_0} se considera de la forma:

$$S_{\beta_0} = \overline{\sigma}\sqrt{\frac{1}{n} + \frac{\overline{x}^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2}}. \tag{2.20}$$

Donde $\overline{\sigma}$ es obtenido por 2.17. El intervalo de confianza para el estimador del intercepto está dado por:

$$IC = \left(\hat{\beta}_0 - t_{(\frac{\alpha}{2},n-2)}S_{\beta_0}, \hat{\beta}_0 + t_{(\frac{\alpha}{2},n-2)}S_{\beta_0}\right). \tag{2.21}$$

Donde n son los grados de libertad y α es el nivel de significancia de la prueba.

2.5. Regresión lineal múltiple

Cuando un fenómeno tiene como respuesta a la interacción de más de una variable es necesario recurrir al modelo de regresión múltiple. En este caso, en lugar de considerar a cada muestra como un valor escalar x, se tiene en su lugar un vector $(x_1, \ldots, x_k)$ para cada punto i de la muestra (Montgomery y col., 2012).

Sea n el número de observaciones, cada una de estas contará con k variables predictoras diferentes. Al predecir y para cada punto de la muestra como una función lineal de las diferentes variables x tenemos como resultado:

$$y = \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k. \tag{2.22}$$

Este última ecuación mantiene su linealidad, por lo que es posible evaluar procedimientos con más de una variable influyente en sus resultados. Sin embargo, podemos usar modelos múltiples para interpretar fenómenos con funciones cuadráticas o polinomiales, tales son los casos de fenómenos con comportamientos parabólicos o con exponentes de mayor grado.

La ecuación 2.23 es un ejemplo de modelos más complejos que pueden ser analizados por técnicas de modelos de regresión múltiple.

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \epsilon. \tag{2.23}$$

En general, cada vez que se requiere conocer la existencia de una relación entre dos o más variables se puede recurrirá al análisis de regresión.

Usar modelos de regresión lineal nos permite descubrir la relación entre estas a través de una profunda investigación de cada una de las variables asi como de un exhaustivo análisis estadístico.

2.5.1. Estimación de los coeficientes de la regresión por mínimos cuadrados

Se representa al modelo como una matriz X con dimensión $n \times k$ donde cada fila corresponde a un punto de la muestra y cada columna a una variable independiente. Desde que cada resultado y_i es un escalar, esta colección de escalares se denotará por el vector columna y con dimensión $n \times 1$.

El modelo lineal será expresado de la manera:

$$y = X\beta + \epsilon.$$

Donde β es un vector de k elementos con dimensión $k \times 1$ y ϵ es una matriz de n elementos con dimensión de $n \times 1$. Cada uno de los ϵ_i es una variable aleatoria con distribución normal con media 0 y varianza constante.

El problema de optimización se define como:

$$\min_{\beta} \sum_{i=1}^{n} (y_i - X_i\beta)^2.$$

Donde se busca encontrar los valores de β que minimizan el problema y Xi son referidos a las filas de la columna X. Usando conocimientos del álgebra lineal se resuelve el problema para encontrar el estimador óptimo de mínimos cuadrados ordinarios el cual se expresa como:

$$\hat{\beta} = (X^T X)^{-1} X^T y.$$

Asegurándose con anticipación que la matriz $(X^T X)^{-1}$ existe. La matriz $(X^T X)^{-1}$ siempre existirá si los regresores son linealmente independientes, esto es, si ninguna columna de la matriz X es combinación lineal de otra.

Por otro lado, el vector de los valores estimados dada una muestra con valores observados y_i está dado por:

$$\hat{y} = X\hat{\beta} = X(X^T X)^{-1} X^T y.$$

La diferencia entre los valores observados y_i y los correspondientes valores ajustados teóricos $\hat{y}_i$ es el residual $e_i = y_i - \hat{y}_i$. Los n residuales pueden ser escritos de forma matricial como:

$$e = y - \hat{y}.$$

El uso de la estimación de fenómenos aleatorios a través del método de mínimos cuadrados ordinarios (MCO) es de gran importancia para nuestro estudio debido a las herramientas que nos provee a la hora de analizar y estimar los resultados obtenidos, además de ofrecernos parámetros que pueden ser estudiados usando metodologías como las pruebas de hipótesis para probar su veracidad y presición, tal

como se presentaron en este capítulo.Además de contar con un procedimiento más sencillo, MCO presenta la oportunidad de obtener los resultados validados por una prueba de bondad de ajuste que muchas veces puede ser representados a través de gráficas de probabilidad.
Pese a existir metodologías similares, la metodología de mínimos cuadrados se prefiere debido al contexto de la investigación y la relación que existe entre las variables a estudiar.

El proceso de crecimiento en las especies se puede estimar y pronosticar a través de distintos mecanismos matemáticos. El estudio de esté a través de la metodología de mínimos cuadrados ordinarios refuerza las proyecciones al validar sus resultados utilizando herramientas estadísticas.

Capítulo 3

Metodología

3.1. Panorama demográfico del estado de Puebla

Los resultados del análisis de los componentes demográficos posicionaron, en el año 2015, al estado de Puebla en la quinta posición a nivel nacional con mayor tamaño poblacional, haciéndolo uno de los estados más relevantes en materia demográfica. Esto permitió dar un panorama de lo que la población demandaba en materia de servicios, salud, seguridad, educación, vivienda y empleo para que de esta manera se optimizara la planeación y desarrollo de programas que eventualmente actuarían en ventaja de la sociedad (INEGI, 2009).

La población estatal poblana ha manifestado los mismos patrones de crecimiento respecto a la población nacional. En 1970, el estado contaba con 2.5 millones de personas, dos décadas después, se registraba un aumento de población con valor de 4.7 millones de habitantes para el año 1995 con una densidad de 48.9 % de hombres y 51.1 % de mujeres. De acuerdo con la información del INEGI, para el 2005 Puebla ya contaba con 5.3 millones de pobladores siguiendo proporciones similares entre hombres y mujeres.

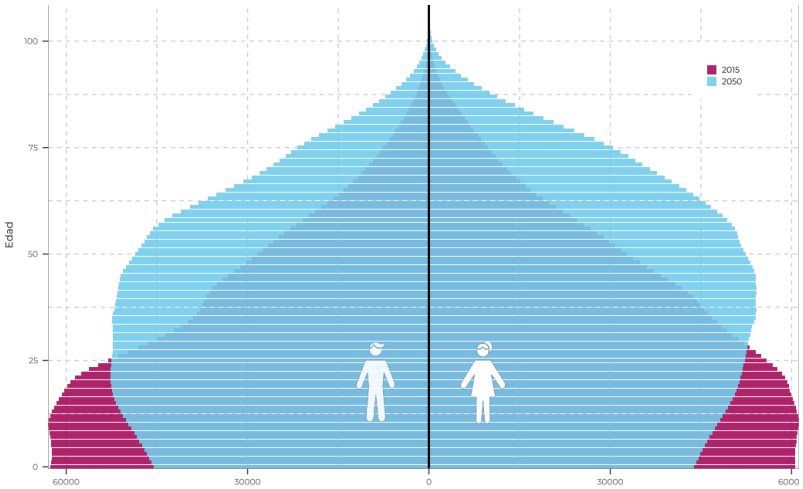
La comparación entre la estructura poblacional del año 2000 y 2005 permite apreciar la reducción de la base de la pirámide demográfica así como el ensanchamiento de los grupos de edad mayores a 30 años, lo que supone una reducción en la tasa de fecundidad estatal. Además durante estos años, el género femenino predomina en el estado resultante de la mayor mortalidad masculina, migración de los varones por asuntos laborales y la tendencia de las mujeres a sobrevivir más en edades adultas (INEGI, 2009).

Instituciones como el Consejo Nacional de Población (CONAPO) establecen un crecimiento de población en las siguientes décadas. Las pirámides poblacionales en los siguientes periodos seguirán mostrando una base amplia, sin embargo, crecerán las poblaciones de personas adultas y mayores de edad.

Una comparación entre las dos pirámides poblacionales de los años 2015 y 2050 realizadas por CONAPO, Figura 3.1, demuestra una evidente reducción de nacimientos. Las jóvenes menores de 15 años solo representarán un 18.6 % de la población para mediados del siglo. Un factor importante recae en las personas en edad pro-

ductiva (15-64 años), tendrán un tamaño considerable 66.7 % en el año 2030 para disminuir en el año 2050. Por otro lado, las personas en edad avanzada seguirán ganando peso en las próximas décadas. En el año 2050 se prevé que representen un 15 % de la población total (CONAPO, 2016).

Figura 3.1: *Población base y estimada para el estado de Puebla, 2015 y 2050.*



Fuente: CONAPO. Proyecciones de la población de México y de las entidades federativas 2016-2050.

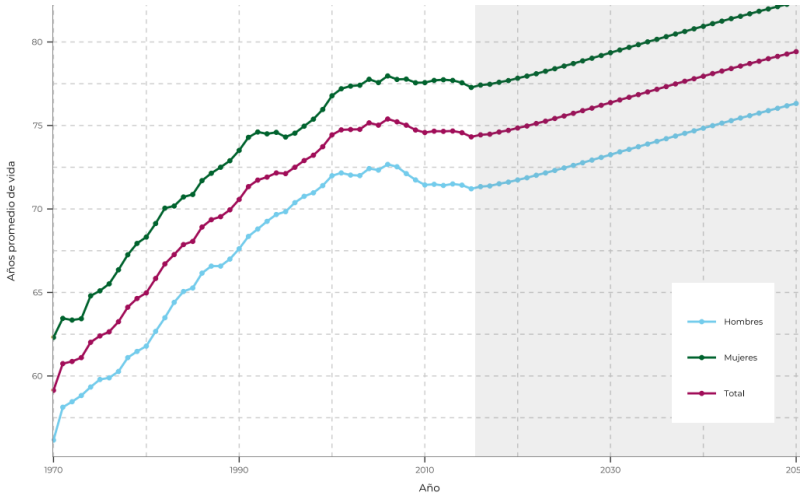
La esperanza de vida en el estado ha presentado avances en las últimas décadas. En un plazo de 35 años, 1970-2005, el tiempo de vida completo aumentó en 15.9 años siendo de 75 años de vida en promedio al final del lapso. A partir de 2005, la tasas de mortalidad, el envejecimiento de la población, el aumento de casos de enfermedades crónico-degenerativas y las muertes debido a situaciones sociales han provocado ligeras fluctuaciones en la variable logrando parar su crecimiento.

La figura 3.2 muestra de manera visual el crecimiento considerable de este fenómeno. En el 2015 la esperanza de vida se registró en 74.3 años, además se estimó una diferencia significante entre la esperanza de vida de hombres y mujeres estimada en 71.2 y 77.3 años promedio.

En contraparte, la Tasa Global de Fecundidad ha preservado su comportamiento descendente en el estado de Puebla. Los factores para explicar su evolución son múltiples, sin embargo, el aumento de la escolaridad, la autonomía de la mujer y la disponibilidad e información sobre los métodos anticonceptivos son los más influyentes. Gracias a esto la planificación familiar y el interés en tener descendencia han reducido el número promedio de hijos de las mujeres durante su vida reproductiva. En el año 1970, la TGF tomaba un valor 7.6, mientras que en 2015, cuarenta y cinco años después, se reducía en 2.28 hijos esperados por cada mujer poblana, véase figura 3.3 (CONAPO, 2016).

La tasa global de fecundidad se pronostica en 1.92 para el 2030 y en 1.71 a mediados del siglo XXI, cuando las mujeres en el rango de 20-24 años conformarán

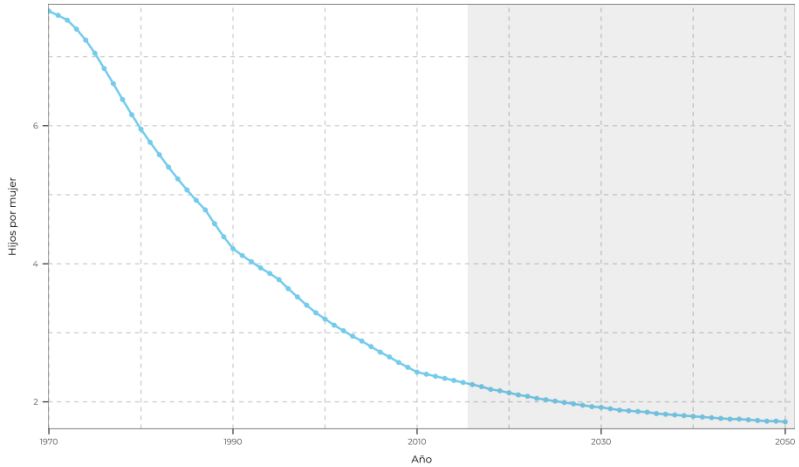
Figura 3.2: *Esperanza de vida media y por sexo para el estado de Puebla, 1970-2050.*



Fuente: CONAPO. Proyecciones de la población de México y de las entidades federativas 2016-2050.

la población de mayor frecuencia del fenómeno reproductivo.

Figura 3.3: *Tasa Global de Fecundidad para el estado de Puebla, 1970-2050.*



Fuente: CONAPO. Proyecciones de la población de México y de las entidades federativas 2016-2050.

3.2. Proyección de la población del Estado de Puebla

Para el presente estudio se emplearon los valores poblacionales del estado de Puebla obtenidos durante los censos poblacionales dirigidos por el Instituto Nacional de Estadística y Geografía (INEGI) comprendidos entre los años de 1895 y 2015, la serie incluye datos censales y de encuestas sociodemográficas, además, de comparaciones con los resultados publicados por el CONAPO mismas que pueden ser encontrados en los portales oficiales de dichas instituciones (CONAPO, 2016), (INEGI, 2015).

A continuación, se obtendrán los resultados emitidos por la investigación bajo las hipótesis de modelos logísticos y la Teoría Estable Acotada para demostrar la existencia de un mínimo y máximo para la media de la población del estado de Puebla del año 1895 a 2015, así como pronosticar el comportamiento demográfico para años posteriores. Las estimaciones, gráficos obtenidos, y los valores críticos de distribuciones fueron realizados y calculados utilizando el software estadístico R versión 3.6.2 (R-STUDIO, 2019).

Para iniciar con el estudio del fenómeno poblacional se muestran las cifras recabadas por INEGI. De manera visual, la tabla 3.1 muestra los resultados poblacionales del estado de Puebla a partir de 1895 hasta el año 2015.

Tabla 3.1: *Crecimiento poblacional del estado de Puebla, 1895-2015.*

AÑO	POBLACIÓN	AÑO	POBLACIÓN
1895	1,025,275	1970	2,508,226
1900	1,021,133	1980	3,347,685
1910	1,101,600	1990	4,126,101
1921	1,024,955	1995	4,624,365
1930	1,150,425	2000	5,076,686
1940	1,294,620	2005	5,383,133
1950	1,625,830	2010	5,779,829
1960	1,973,837	2015	6,168,883

Fuente: INEGI. Censos y Encuestas Sociodemográficas 1895-2015.

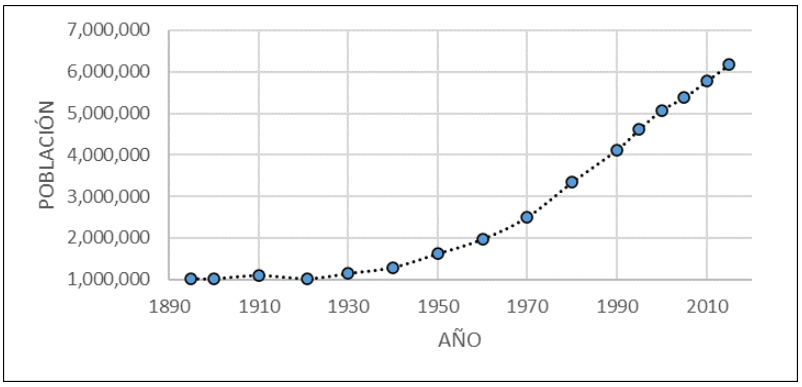
La figura 3.4 muestra un comportamiento creciente de la población con tendencia a seguir aumentando, por lo que se puede plantear que la hipótesis de crecimiento exponencial se presenta en los datos muestrales.

Se observa además que, en el siglo pasado la población se duplicó dos veces. El primer tiempo duplicado fue de 60 años al pasar de poco más de un millón en 1900 a aproximadamente 2 millones en 1960. El segundo tiempo de duplicado sucedió en los siguientes 30 años , lo que indica una desaceleración de crecimiento respecto al tiempo. La evolución observada de la población plantea las siguientes preguntas: ¿habrá un tercer tiempo de duplicación de la población?, ¿existe una cota superior

para el crecimiento de la población?, ¿el patrón de crecimiento de la población puede ser catalogado como logístico ? La respuesta a estas preguntas parecen estar en la Teoría Estable Acotada.

Gráficamente, la figura 3.4 presenta en escala temporal los valores tabulados en la tabla 3.1. Se espera que la población siga la misma distribución y continúe en aumento en el horizonte de tiempo.

Figura 3.4: *Crecimiento poblacional del estado de Puebla 1895-2015.*



Fuente: Elaboración propia en base a la Tabla 3.1

La teoría Estable Acotada se fundamenta en dos importantes postulados (González-Rosas & Zárate-Gutiérrez, 2018).

1. La población es un fenómeno aleatorio, por lo que, de acuerdo con la teoría de la probabilidad, en cada año la población tiene una media y una varianza.

2. La población media está determinada por una función matemática que depende del tiempo, lo cual implica que en cada año la población será explicada por una función matemática más una variable aleatoria. Medhi los nombró el componente determinístico y estocástico respectivamente (Medhi, 1981).

Es decir, bajo estos postulados, el comportamiento de las observaciones y la media de la población en cada lapso serán:

$$P_t = f(t) + \epsilon_t. \tag{3.1}$$

$$\mu_P^t = f(t). \tag{3.2}$$

Donde:

- P_t, denota la población en el tiempo $t \, \epsilon \, R^+$.
- $f(t)$, es una función matemática desconocida.
- μ_P^t, denota la media de la población en el tiempo $t \, \epsilon \, R^+$.

- ϵ_t son errores independientes con distribución normal (gaussiana) con media 0 y varianza constante. Es decir:

$$\epsilon_t \sim \mathcal{N}(0, \sigma^2).$$

Para la obtención del mínimo y máximo de la demografía estatal, la Teoría Estable Acotada usa el incremento de la población a través del tiempo. Dado el incremento entre dos puntos del tiempo, este puede ser medido a través de la pendiente de una línea recta que une dos puntos consecutivos en el espacio bidimensional definido por el tiempo y la población.

La Teoría Estable Acotada prueba que, para dos puntos consecutivos, existen tres estimadores asociados tanto al mínimo como al máximo de la población. Uno para el menor de los valores, otro para el valor mayor y uno más para el valor medio. La Teoría Estable Acotada prueba además que el mejor estimador es el asociado al valor medio. Definimos la pendiente de la línea recta y los puntos medios entre dos lapsos consecutivos de tiempo de la siguiente manera:

$$\nabla_t^P = \frac{P_{t_{i+1}} - P_{t_i}}{t_{i+1} - t_i}. \tag{3.3}$$

$$PM_t = \frac{P_{t_{i+1}} + P_{t_i}}{2}. \tag{3.4}$$

Donde:

- ∇_t^P,denota la pendiente de la línea recta que une a los puntos (P_{t_i}, t_i) y $(P_{t_{i+1}}, t_{i+1})$ del espacio bidimensional tiempo-población.

- PM_t, representa el valor medio entre los valores $(P_{t_i}, P_{t_{i+1}})$.

El gran aporte de la Teoría Estable Acotada es suponer que existe una relación entre la pendiente de los puntos consecutivos de la población y la misma población, cambiando el análisis de un espacio bidimensional tiempo-población a un espacio bidimensional pendiente-población.

La tabla 3.2 presenta los valores obtenidos dados en millones para las poblaciones medias entre cada pareja de puntos consecutivos de la muestra, así como para las pendientes de cada línea recta entre los puntos de la misma. La figura 3.5 representa los puntos medios en el eje X y el valor de las pendientes en el eje Y. En la figura 3.5 el comportamiento de las pendientes con respecto de los valores poblaciones parecen seguir un patrón parabólico. Si esta hipótesis es cierta, de acuerdo con el patrón, se observa que la pendiente fue cero en el pasado muy cerca del valor poblacional de un millón. Después sigue en aumento a medida que se incrementa la población hasta que alcanza el máximo aproximadamente en los 4.5 millones de habitantes. Una vez alcanzado el máximo, la pendiente empezó a descender y se prevé que en el futuro volverá a ser cero muy cerca de los ocho millones de habitantes. Estos resultados prueban empíricamente el supuesto de la Teoría Estable Acotada.

Tabla 3.2: *Tabulación de los puntos medios y pendientes 1895-2015.*

AÑO	TIEMPO	POBLACIÓN	PM_t	∇_t^P
1895	0	1.025275	1.0232040	-0.000828
1900	5	1.021133	1.0613665	0.008046
1910	15	1.101600	1.0632775	-0.006967
1921	26	1.024955	1.0876900	0.013941
1930	35	1.150425	1.2225225	0.014419
1940	45	1.294620	1.4602250	0.033121
1950	55	1.625830	1.7998335	0.034800
1960	65	1.973837	2.2410315	0.053438
1970	75	2.508226	2.9279555	0.083945
1980	85	3.347685	3.7368930	0.077841
1990	95	4.126101	4.3752330	0.099652
1995	100	4.624365	4.8505255	0.090464
2005	110	5.383133	5.5814810	0.079339
2010	115	5.779829	5.9743560	0.077810
2015	120	6.168883		

Fuente: Cálculos basados en las ecuaciones 3.3 y 3.4. El año 2000 fue omitido por considerarse dato atípico.

El movimiento parabólico mostrado en la figura 3.5 puede ser considerado un polinomio de segundo grado.

Se puede demostrar que, un polinomio de segundo grado (línea punteada de la figura 3.5), debe tener dos raíces (Beaumont & Pierce, 1963), es decir, dos puntos de intersección entre la curva y el eje X. Como el concepto graficado es la pendiente, entonces en la intersección, el grado de inclinación es cero y por lo tanto es paralelo al eje del tiempo t, lo que implica que los valores donde esta toma el valor 0, pues es donde el mínimo y el máximo se presentan.

Llamaremos a la primera raíz K y a la segunda raíz $K + C$. Donde K será el mínimo de la población estatal, mientras que el valor $K + C$ será el máximo de la misma.

Estos serán los valores por los que el valor estimado para el crecimiento demográfico del estado de Puebla se encontrará acotado a través del tiempo.

En primera instancia, se debe demostrar su existencia, para ello, ajustamos el siguiente modelo de regresión lineal múltiple a los datos de la figura 3.5. Se define el siguiente modelo:

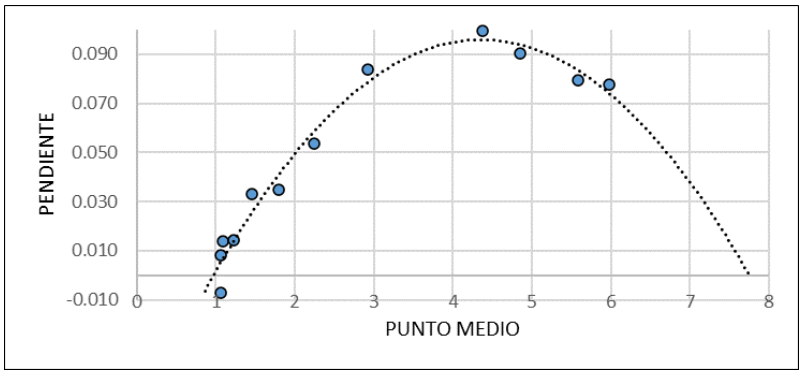
$$\nabla_{t_i}^P = APM_{t_i}^2 + BPM_{t_i} + D + \pi_i. \tag{3.5}$$

$$\mu_{\nabla}^P = APM_{t_i}^2 + BPM_{t_i} + D. \tag{3.6}$$

Donde:

- A, B, D, son constantes desconocidas.

Figura 3.5: *Gráfica de los puntos medios y pendientes 1895-2015.*



Fuente: Elaboración propia en base a la Tabla 3.2 y bajo las relaciones 3.3 y 3.4.

- PM_{t_i}, denota el punto medio entre dos valores consecutivos de la población.
- $\nabla^P_{t_i}$, es el valor de la pendiente en el tiempo t_i.
- π_i, son los errores independientes con distribución normal $\mathcal{N}(0, h)$, donde h es una constante.
- μ^P_{∇}, es la media de la variable aleatoria $\nabla^P_{t_i}$.

Para estimar los valores de A, B, D se considera el modelo de regresión lineal múltiple dado en 2.22, donde los valores de las pendientes se consideran como variable dependiente, mientras que el valor medio de la población y su valor cuadrático componen el campo de las variables explicativas.

Los resultados de la estimación bajo la metodología de Mínimos Cuadrados Ordinarios son presentados en la tabla 3.3.

Tabla 3.3: *Parámetros obtenidos del modelo dado por la ecuación 3.6.*

Parámetro	Estimación	Error estándar	Valor T_0	Valor P
A	-0.007699	0.000935	-8.233	$4.97*10^{-6}$
B	0.068381	0.006221	10.99	$2.86*10^{-7}$
D	-0.058888	0.007785	-7.564	$1.11*10^{-5}$
Coeficiente de determinación ajustado (R^2)			Valor F_0	Valor P
0.9613			162.3	$6.84*10^{-9}$

En la tabla 3.3, se puede ver la significancia individual de los parámetros a través de su respectiva Prueba T (Apéndice B), o bien, el Valor P (Apéndice E). Recordando que la hipótesis nula, (considerando a $\beta_i = 0$) determina que el estimador no es significante para la prueba. Rechazamos H_0 si $T_0 \leq -t_{\frac{\alpha}{2},n-k-1}$ ó $T_0 \geq t_{\frac{\alpha}{2},n-k-1}$ o se cumple que *Valor* $P \leq \alpha$. Con una muestra de tamaño 14 y un nivel de significancia del 5 %, el valor crítico de la distribución T toma el valor $t_{0.025,11} = 2.2$. En

la Tabla 3.3 se muestran los parámetros con sus respectivos valores estadísticos T_0. Se observa que todos los estadísticos caen dentro de la zona de rechazo, además de contar con valores P menores al 5 %, por lo que es evidente rechazar la hipótesis nula. Esto prueba que las constantes son significativamente diferentes de cero.

Por su parte, para la estadística F (Apéndice C) que prueba la significancia total del modelo, obtenemos un estadístico de prueba de $F_0 = 162.3$, contra un valor crítico $F^2_{11,0.05} = 3.98$, por lo que rechazamos también la hipótesis nula de que todos los constantes son iguales a cero y se inclina por la significancia global del modelo además de obtener un valor P para la prueba F menor a 0.05. Además, el coeficiente de determinación R^2 (Apéndice D) muestra que el modelo explica el 96.13 % de la variación total de la variable dependiente. Estos resultados son una prueba matemática de que el patrón parabólico es cierto.

De acuerdo con la Teoría Estable Acotada, desde el punto de vista geométrico, el máximo y el mínimo de la población poblana están dados por los valores en los que la curva de la gráfica 3.5 intersecta el eje de las X, pero desde el punto de vista matemático el máximo y el mínimo son las raíces del polinomio de segundo grado dado en 3.6. Como es sabido, el máximo y el mínimo ocurren cuando el valor de la pendiente es igual a 0. Es decir:

$$APM_{t_i}^2 + BPM_{t_i} + D = 0. \tag{3.7}$$

Usando la fórmula general para la obtención de raíces reales de un polinomio de segundo grado, obtenemos que:

$$K = \frac{-B + \sqrt{B^2 - 4AD}}{2A}. \tag{3.8}$$

$$K + C = K = \frac{-B - \sqrt{B^2 - 4AD}}{2A}. \tag{3.9}$$

Estos resultados indican que las relaciones 3.8 y 3.9 son estimadores para el valor mínimo y máximo de la población respectivamente, los cuales existen siempre que $B^2 - 4AD \geq 0$.

Utilizando las estimaciones de los coeficientes en las ecuaciones 3.8 y 3.9 se obtienen las cotas poblacionales. Estas son obtenidas como sigue:

$$K = \frac{-\hat{B} + \sqrt{\hat{B}^2 - 4\hat{A}\hat{D}}}{2\hat{A}} = \frac{-0.0683 + \sqrt{(0.0683)^2 - 4(-0.00769)(-0.0588)}}{2(-0.00769)}$$

$$K = 0.966315$$

$$K + C = \frac{-\hat{B} - \sqrt{\hat{B}^2 - 4\hat{A}\hat{D}}}{2\hat{A}} = \frac{-0.0683 - \sqrt{(0.0683)^2 - 4(-0.00769)(-0.0588)}}{2(-0.00769)}$$

$$K + C = 7.915555$$

Con un nivel de significancia del 5 %, valores P menores a 0.05 y estadísticos altamente significantes al 95 % de confianza se prueba la existencia del mínimo y máximo de la media de la población para el estado de Puebla. Cabe aclarar, que estas cotas son para la media de la población, más no para las observaciones ya que, por la teoría de la probabilidad, estas siempre estarán desviadas de la media tanto como lo permita su varianza, Es decir, las observaciones pueden tomar valores más pequeños que el mínimo y mayores al máximo, sin embargo, su distribución siempre estará regida por una ley de probabilidad.

Los valores para dichas cotas están dados por las raíces $K = 0.9566315$ y $K + C = 7.915555$. Estos datos comprueban lo apreciado en la figura 3.5.

De acuerdo con la Teoría Estable Acotada, el comportamiento de la demografía estatal en cada tiempo se determina por 3.1 la cual incluye una función desconocida. Sin embargo, bajo el contexto del fenómeno demográfico, está puede ser estimada.

Como se puede observar en los datos recabados en la Tabla 3.1 y en la gráfica 3.4, la población muestral sigue una función creciente a través del tiempo, por lo que la derivada de dicha función debe ser positiva. Además, la función evaluada en el mínimo y máximo deben tomar el valor 0.

Para estimar la función, la Teoría Estable Acotada supone que la derivada de $f(t)$ está dada por:

$$\frac{\mathrm{d}P}{\mathrm{d}t} = h(P)j(t) \tag{3.10}$$

Donde $\frac{\mathrm{d}P}{\mathrm{d}t}$, denota la derivada de $f(t)$.

Es decir, la derivada de la función desconocida dada en 3.1 puede ser vista como el producto de dos funciones, la primera estará en función de la población, mientras que la segunda estará en función del tiempo.

Ahora, por el comportamiento de los datos y por propiedades del máximo y mínimo, la derivada $\frac{\mathrm{d}P}{\mathrm{d}t}$, debes ser positiva y ser igual a 0 en los valores K y $K + C$. Por lo anterior, la función determinada por la población puede ser vista como:

$$h(P) = (P - K)(P - K - C). \tag{3.11}$$

Ya que si se evalúa la función en el mínimo y máximo, la función $h(P)$ se anula.

$$h(K) = (K - K)(K - K - C) = 0.$$

$$h(K + C) = (K + C - K)(K + C - K - C) = 0.$$

Y por consecuencia la derivada también se anula en esos mismos puntos. Es decir:

$$\frac{\mathrm{d}P}{\mathrm{d}t}(K) = h(K)j(t) = 0.$$

$$\frac{\mathrm{d}P}{\mathrm{d}t}(K + C) = h(K + C)j(t) = 0.$$

Sustituyendo 3.11 en 3.10, la derivada de la función desconocida puede ser expresada como:

$$\frac{\mathrm{d}P}{\mathrm{d}t} = (P-K)(P-K-C)j(t). \tag{3.12}$$

En donde se observa que al ser K el mínimo esperado de la población, $(P-K)$ será siempre positiva. Por otro lado, al ser $K+C$ el máximo esperado, $(P-K-C)$ será siempre negativo, por lo que el producto $(P-K)(P-K-C)$ debe poseer un signo negativo. Se ha dicho que al ser la población un fenómeno creciente, su derivada $\frac{\mathrm{d}P}{\mathrm{d}t}$ debe tener un signo positivo. Por lo tanto, para que la función dada por 3.12 sea positiva, es necesario que la función que depende del tiempo $j(t)$ tenga signo negativo.

La relación 3.10 expresa una ecuación diferencial ordinaria de primer orden de variables separables que guarda relación con la función logística explicada en capítulos anteriores. Usando el procedimiento sistematizado del método de variables separables (Bravo, 2017) se obtiene la relación:

$$\int \frac{1}{(P-K)(P-K-C)}\,\mathrm{d}P = \int j(t)\,\mathrm{d}t. \tag{3.13}$$

Como la función que depende del tiempo es desconocida, el valor de su integral también lo es. Por lo que se establece la siguiente igualdad:

$$\lambda(t) = \int j(t)\,\mathrm{d}t.$$

Donde, $\lambda(t)$ es una función desconocida tal que su derivada es igual a $j(t)$.

Resolviendo por el método de integración de fracciones parciales (Del-Pino, 2019), se obtiene la solución de la ecuación diferencial y despejando la variable P se obtiene la expresión para la población (ver Apéndice A). La solución queda expresada como:

$$P = K + \frac{C}{1+e^{\,c\lambda(t)}}. \tag{3.14}$$

De 3.14 se puede obtener una estimación para la función $\lambda(t)$ en base a una muestra dada.

La función puede ser determinada por:

$$\lambda(t) = \frac{1}{C}\ln\left(\frac{C}{P-K}-1\right). \tag{3.15}$$

Lo que implica directamente que si la Teoría Estable Acotada es cierta, la variable del lado derecho de 3.15 debe estar en función del tiempo. Esta variable se conoce como la *transformada de la población* (González-Rosas & Zárate-Gutiérrez,

2018).

Para la estimación de la transformada de la población $\lambda(t)$ se usa que:

$$K = 0.9663151$$

$$K + C = 7.915555$$

$$C = 6.949239$$

Con los resultados anteriores y los valores observados para la población estatal obtenidos en la muestra, se obtiene la Tabla 3.4. En esta, se emiten los valores obtenidos para la transformada de la población a través de la línea temporal de la muestra. Se puede observar en la figura 3.6 como la transformada de la población presenta

Tabla 3.4: *Función transformada 1895-2015.*

AÑO	TIEMPO	POBLACIÓN	$\lambda(t)$
1895	0	1.025275	0.6851
1900	5	1.021133	0.6957
1910	15	1.101600	0.5640
1930	35	1.150425	0.5186
1940	45	1.294620	0.4323
1950	55	1.625830	0.3245
1960	65	1.973837	0.2554
1970	75	2.508226	0.1806
1980	85	3.347685	0.0937
1990	95	4.126101	0.0261
1995	100	4.624365	-0.0152
2000	105	5.076686	-0.0533
2005	110	5.383133	-0.0800
2010	115	5.779829	-0.1169
2015	120	6.168883	-0.1571

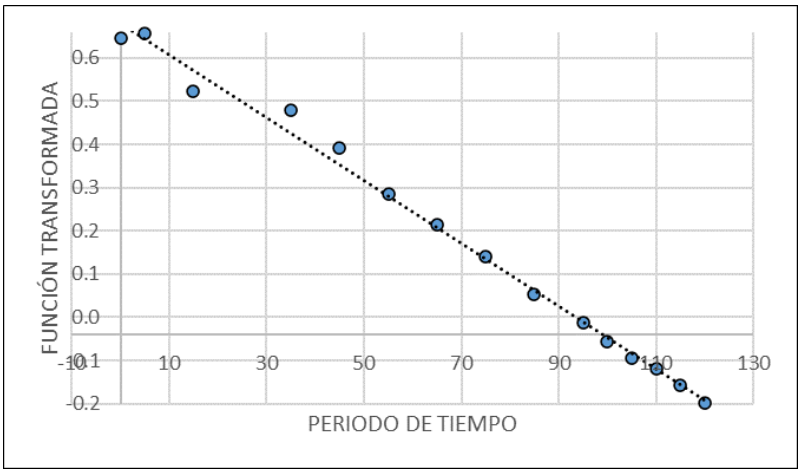
Fuente: Cálculos basados en la ecuación 3.15. El dato del año 1921 fue omitido por considerarse dato atípico. Ningún dato fue omitido por indeterminar la función logaritmo por lo que está definida en cada punto de la muestra.

un comportamiento lineal a través del tiempo. Si esta hipótesis es correcta, implica que la función transformada posee la forma:

$$\lambda(t) = \alpha \mathrm{t} + \beta. \tag{3.16}$$

Volviendo a la naturaleza y contexto del problema, la derivada de la función $\lambda(t)$ debe ser igual a la función respecto al tiempo $j(t)$, la cual debe tener signo negativo. Esto tiene como consecuencia que la función $\lambda(t)$ debe ser decreciente respecto al tiempo ya que su derivada posee signo negativo.

Figura 3.6: *Gráfica de la función transformada respecto al tiempo, 1895-2015.*



Fuente: Elaboración propia en base a la Tabla 3.4 y la ecuación 3.15

En otras palabras, si la función sigue un patrón lineal $\lambda(t) = \alpha t + \beta$ y la función $j(t)$ tiene signo negativo, el parámetro de la pendiente de la recta α debe ser negativo también.

Si el modelo lineal es aceptado se obtiene entonces un patrón logístico para la población, puesto que si la estimación del modelo se sustituye en 3.14, la ecuación para dicho patrón corresponde a:

$$P = K + \frac{C}{1 + e^{c(\alpha t + \beta)}} \tag{3.17}$$

Respecto a la ecuación 3.17, usando el supuesto que $\alpha < 0$, cuando $t \to \infty$, el valor de la población del estado de Puebla se aproxima a la cota $K + C$, de manera formal se dice que $\lim_{t \to \infty} P_t = K + C$.

La velocidad de que tan rápido la población converge al máximo depende de la magnitud de los parámetros α y β, por esta razón, estos estimadores reciben el nombre de *parámetros de rapidez* (González-Rosas & Zárate-Gutiérrez, 2018).

Para estimar los parámetros de rapidez se realiza un modelo de regresión lineal simple bajo la metodología de MCO tomando como variable independiente el tiempo, mientras que los valores obtenidos para $\lambda(t)$ son considerados como la variable dependiente, ambas variables se consideran en la tabla 3.4.

La tabla 3.5 muestra las estimaciones para los parámetros de la recta ajustada al modelo correspondiente a la figura 3.6, además de presentar los valores de los estadísticos T_0 y F_0, así como los respectivos valores P de las distintas pruebas estadísticas.

Se observa que, para ambos parámetros, la estimación es significativa a un nivel

Tabla 3.5: *Estimadores para los parámetros de rapidez.*

Parámetro	Estimación	Error estándar	Valor T_0	Valor P
α	-0.0072594	0.0001	-43.84	$1.64 * 10^{-15}$
β	0.7196245	0.0130	55.09	$2 * 10^{-16}$
Coeficiente de determinación ajustado (R^2)			Valor F_0	Valor P
0.9928			1922	$1.64 * 10^{-15}$

de confianza del 95 %. Se cuenta con un estadístico T_0 con magnitud de -43.84 para la estimación de la pendiente, mientras que el valor T_0 del intercepto estimado es de 55.09.

Con una muestra de tamaño $n = 15$, el valor crítico de la distribución t toma el valor $t_{(\frac{0.05}{2},13)} = 2.16$, por lo tanto, ambos estimadores caen en la región de rechazo. Se respalda que los valores P son menores del 0.05, esto implica que su aportación al modelo es significante ya que se rechaza la hipótesis nula. Utilizando la Prueba F para probar la significancia total del modelo obtenemos un estadístico de prueba de $F_0 = 1922$, contra un valor crítico $F^1_{13,0.05} = 4.66$, una vez más, rechazamos la hipótesis nula y se inclina por la significancia global del modelo además de obtener un valor P para la prueba F menor a 0.05. Con un nivel del 99.28 % de variación explicada por el estadístico R^2 ajustado, es evidente la significancia del modelo.

Con todas las pruebas estadísticas a favor del modelo, se concluye que la transformada de la población sigue un patrón lineal, además de ser decreciente respecto al tiempo, lo que prueba las predicciones de la Teoría Estable Acotada.

Se afirma entonces que la función se determina por la relación:

$$\lambda(t) = -0.0072594t + 0.7196245. \tag{3.18}$$

Al sustituir 3.18 en la función 3.17 se obtiene el modelo logístico poblacional para el estado de Puebla bajo la metodología de modelos logísticos, estimaciones por la metodología MCO, en base a los postulados en 3.2 y dada la muestra en la tabla 4.1.

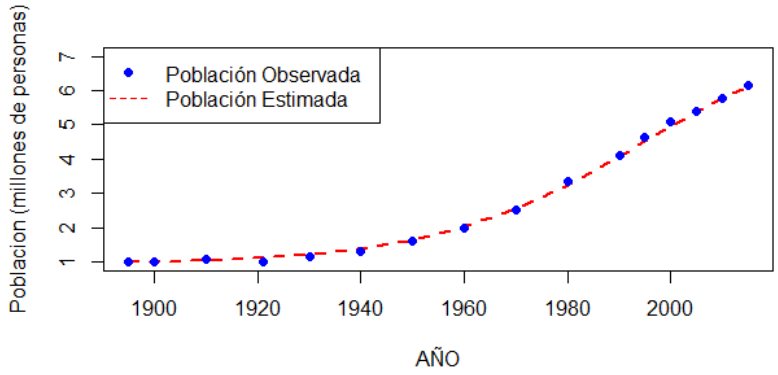
Dicha estimación se resume por la expresión:

$$P_t = 0.9663151 + \frac{6.949239}{1 + e^{6.949239(-0.0072594t + 0.7196245)}} \tag{3.19}$$

Donde P_t es la media de la población en el tiempo t.
En la figura 3.7 se muestra el modelo logístico obtenido contrastado con los puntos de la muestra inicial. Se observan valores estimados con gran persistencia sobre los valores poblacionales observados.

En la tabla 3.6 se puede apreciar la evolución de la demografía de manera creciente en la mayoría de los datos de la muestra, a excepción de los años 1900 y 1921 donde el fenómeno resultó contrario a lo esperado puesto que las diferencias se deben

Figura 3.7: *Comparación entre la población estatal observada y estimada de Puebla.*



Fuente: Elaboración propia en base a la Tabla 3.1 y los resultados de la ecuación 3.19.

a que la ecuación predice la media y no las observaciones. Por otro lado, la población estimada sigue un comportamiento estrictamente creciente pues el modelo dado por la relación 3.19 lo describe de esa manera. Se muestra una evaluación aceptable respecto a lo observado y se espera que la población continúe bajo la misma evolución.

Se considera al error relativo que se produce durante el proceso de estimación obtenida de la sustracción entre la población observada y la estimada. Se observan magnitudes pequeñas por lo que las estimaciones son aceptadas.

El último objetivo es obtener las proyecciones poblacionales del estado para años posteriores a 2015. A partir de la relación 3.19 y utilizando el software estadístico se calculan los pronósticos. La tabla 3.7 detalla las medias poblacionales pronosticadas para el periodo comprendido entre los años 2020-2050 así como los intervalos al 80 % y 95 % de confianza.

Dichos intervalos contendrán el valor desconocido de la población en 80 % y 95 % de confianza de acierto respectivamente, por lo tanto, para fines de análisis se considera el intervalo al 95 % puesto que existe mayor confianza de que el valor futuro exacto se encuentre comprendido en él. Cabe recalcar que las estimaciones están basadas en las medias poblacionales en cada periodo del tiempo, por lo tanto existirá un error aleatorio en cada lapso. El tamaño de estos errores dependerá de la varianza obtenida del modelo de predicción.

La figura 3.8 contrasta la población estudiada junto con los pronósticos realizados hasta el año 2100. Esto comprueba la hipótesis de la existencia del patrón logístico en la población.
Además, se puede apreciar el comportamiento convergente sobre el intervalo de con-

Tabla 3.6: *Población observada, estimada y residual para los años 1895-2015. Unidades en millones de personas.*

Año	Población observada	Población estimada	Error relativo
1895	1.025275	1.012786	0.012181
1900	1.021133	1.026004	0.004770
1910	1.101600	1.064613	0.033576
1921	1.024955	1.135753	0.108100
1930	1.150425	1.229432	0.068676
1940	1.294620	1.391503	0.074835
1950	1.625830	1.643297	0.010743
1960	1.973837	2.020117	0.023447
1970	2.508226	2.553601	0.018090
1980	3.347685	3.252432	0.028453
1990	4.126101	4.080256	0.011111
1995	4.624365	4.517152	0.023184
2000	5.076686	4.951650	0.024629
2005	5.383133	5.370414	0.002363
2010	5.779829	5.761984	0.003087
2015	6.168883	6.117901	0.008264

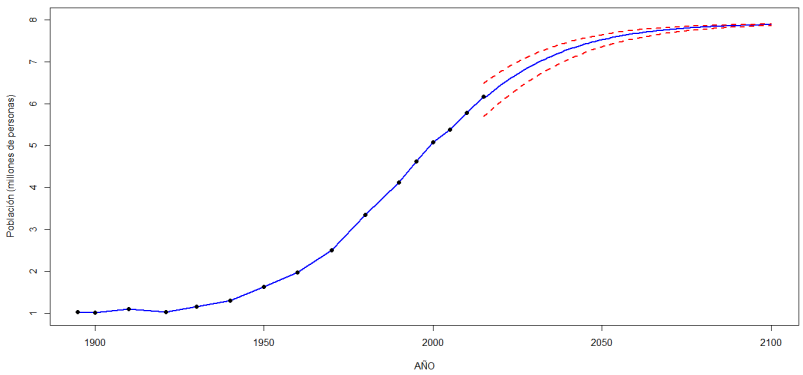
fianza de las estimaciones. Se observa de manera clara la disminución del tamaño del intervalo a medida que se recorre en el tiempo, esto se le puede atribuir al hecho de que la población converge a una constante, que de hecho es su máximo, por lo que el intervalo de incertidumbre cada vez tendrá que reducirse pues dicha incertidumbre se ve reducida con el paso del tiempo.

Se muestra un comportamiento logístico ascendente ligeramente lineal. Además, por la Teoría Estable Acotada se espera que cada año la tasa de crecimiento se reduzca año con año hasta presentar un comportamiento casi lineal cuando se alcance la estabilidad. Por otra parte, se observan intervalos de confianza (líneas punteadas) con una longitud considerable al inicio de la proyección. Puede notarse ligeramente la disminución sobre el tamaño del intervalo respecto al avance del tiempo, esto puede ocurrir debido a que la tendencia de los pronósticos se aproxima a una constante conforme el tiempo aumenta. Cuando esto ocurra, la población estatal convergerá a su máximo y se estabilizará.

Se espera que al final de la tercera década del siglo XXI, la media poblacional del estado de Puebla sea próxima de 7 millones de personas, después de esto, bajo la Teoría Estable Acotada, la capacidad del entorno solo podrá soportar, aproximadamente, a un millón de habitantes más a partir del año 2030.

Es de suma importancia que las autoridades tanto estatales como federales asimilen la evolución temporal de la población y que se mantengan atentas en las repercusiones sociales, económicas, ambientales y legislativas que pueden ramificar-

Figura 3.8: *Población observada en el lapso 1895-2015 y pronósticos para la media poblacional con intervalos de confianza al 95 % 2020-2100.*



Fuente: Elaboración propia en base a los resultados presentes en las Tablas 3.1 y 3.7.

se debido a la movilización y reestructuración de la pirámide poblacional. Eventos como la disminución de la tasa global de fecundidad y aumento en la esperanza de vida supondrían un envejecimiento paulatino de la población.

A medio y largo plazo estos acontecimientos se evidenciarán en el contexto social, por lo que las autoridades competentes comenzarían a crear reformas en los servicios educativos y todos los relacionados con la salud tales como el mejoramiento de los programas de asistencia para los grupos de personas de la tercera edad y de grupos vulnerables, además de contemplar nuevas opciones para el ampliamiento en la cobertura en el sistema de pensiones a nivel nacional debido al aumento en la población en edad de retiro a causa del envejecimiento de la población (Villa-D, 2019).

Tabla 3.7: *Pronósticos para la media poblacional del estado de Puebla, 2020-2050.*

AÑO	Pronóstico	L. Inf 80 %	L. Sup 80 %	L. Inf 95 %	L. Sup 95 %
2020	6,433,178	6,105,068	6,715,709	5,886,951	6,863,765
2021	6,491,157	6,171,067	6,765,492	5,957,407	6,908,765
2022	6,547,436	6,235,437	6,813,613	6,026,339	6,952,165
2023	6,602,019	6,298,159	6,860,094	6,093,712	6,993,995
2024	6,654,917	6,359,222	6,904,960	6,159,500	7,034,286
2025	6,706,142	6,418,618	6,948,238	6,223,680	7,073,071
2026	6,755,710	6,476,343	6,989,958	6,286,234	7,110,384
2027	6,803,640	6,532,398	7,030,151	6,347,151	7,146,261
2028	6,849,953	6,586,788	7,068,850	6,406,421	7,180,739
2029	6,894,675	6,639,521	7,106,089	6,464,042	7,213,855
2030	6,937,832	6,690,611	7,141,902	6,520,014	7,245,647
2031	6,979,465	6,794,729	7,137,816	6,670,352	7,221,185
2032	7,019,580	6,839,934	7,173,015	6,718,621	7,253,574
2033	7,058,221	6,883,639	7,206,804	6,765,396	7,284,605
2034	7,095,421	6,925,867	7,239,223	6,810,694	7,314,321
2035	7,131,214	6,966,644	7,270,313	6,854,534	7,342,764
2036	7,165,637	7,00,5998	7,300,113	6,896,938	7,369,977
2037	7,198,725	7,043,956	7,328,666	6,937,929	7,396,003
2038	7,230,514	7,080,550	7,356,010	6,977,531	7,420,883
2039	7,261,041	7,115,809	7,382,188	7,015,771	7,444,659
2040	7,290,344	7,149,766	7,407,238	7,052,675	7,467,372
2041	7,318,460	7,182,453	7,431,201	7,088,273	7,489,061
2042	7,345,425	7,213,903	7,454,115	7,122,594	7,509,766
2043	7,371,276	7,244,149	7,476,018	7,155,667	7,529,525
2044	7,396,051	7,273,226	7,496,949	7,187,525	7,548,376
2045	7,419,786	7,301,167	7,516,944	7,218,197	7,566,356
2046	7,442,516	7,328,006	7,536,040	7,247,715	7,583,500
2047	7,464,277	7,353,776	7,554,272	7,276,112	7,599,842
2048	7,485,103	7,378,511	7,571,673	7,303,420	7,615,416
2049	7,505,029	7,402,245	7,588,278	7,329,669	7,630,256
2050	7,524,087	7,425,010	7,604,119	7,354,892	7,644,391

Fuente: Cálculos de los pronósticos basados en la relación 3.19.

Capítulo 4

Análisis y conclusiones

En base a todas las evidencias recabadas y usando metodologías estadísticas, se demuestra que la población humana puede estimarse con el uso de modelos logísticos, pues concibe sus propias restricciones en su capacidad de entorno, mismas que pueden ser de índoles económicas, sociales, educativas, salubres, ambientales etc.

La relación población-tiempo ha sido explicada tomando en cuenta la media poblacional en cada punto de la muestra y utilizando su variabilidad respecto al siguiente periodo de la misma. Los datos presentes en dicha muestra están completamente respaldados por el Instituto Nacional de Estadística y Geografía a través de los censos poblacionales celebrados cada diez años y sus encuestas intercensales realizadas cada cinco.

Para llevar a cabo el análisis demográfico se realizó un estudio de los datos de manera descriptiva, es decir se hizo uso de gráficos para observar el comportamiento tanto de la población general como de la evolución de las pendientes respecto al tiempo y medias poblaciones, respectivamente. Para el primer gráfico (figura 3.1), se detectó un crecimiento exponencial en su crecimiento entre los años 1895 a 1990, sin embargo, a partir de este año hasta el 2015 se comprobó que la evolución del fenómeno demográfico resulta en un modelo logístico.

Para el segundo gráfico (figura 3.2), se obtuvo una curva parabólica, de la cual, sus raíces o intercesiones con el eje horizontal resultaron en las cotas poblacionales presentadas en este trabajo. Esto resulta debido a la disminución en la tasa de crecimiento poblacional a través del periodo de tiempo comprendido por la muestra. Puede observarse que los resultados respaldan lo demostrado por la primera gráfica, pues entre los años ochenta y noventa la población presentó un crecimiento exponencial, debido a que durante este lapso la tasa de crecimiento se desarrolló de manera crecimiento, llegando a su máximo en el año de 1990. A partir de entonces, la curva poblacional mantiene su comportamiento bajo una curva logística.

Una vez demostrado su comportamiento se estimó el modelo de segundo grado a partir de la metodología de mínimos cuadrados ordinarios, resultando en conclusiones ya esperadas sometidas a diferentes pruebas estadísticas que avalan el comportamiento de un polinomio parabólico. Respecto a la estimación de la curva se decidió

omitir los resultados del año 2000 en la Tabla 3.2 debido a que una aproximación con este dato dentro de la muestra resultaba en una curva más cerrada. Esto provocaba un aumento considerable en la cota mínima poblacional y una disminución por más de 200,000 personas para la cota máxima, además de restar significancia a las pruebas T y F para los estimadores de la curva. Por lo tanto, se consideró un dato atípico.

Después de estimar un modelo con una significancia estadística considerable, se le asignan los nombres de cota mínima y máxima poblacional a las raíces reales del polinomio obtenido bajo la metodología de MCO. Se espera que, con el avance del tiempo, la población converja asintóticamente al máximo de la demografía estatal.

Utilizando el método de integración de fracciones separables y el proceso sistematizado de ecuaciones diferenciales de varias separables, se obtuvo una estimación para la función transformada de la población. Cumpliendo con la teoría, se demostró su comportamiento decreciente respecto al tiempo y bajo la metodología de MCO, se obtuvo una tendencia lineal (figura 3.3).

Para su aproximación, se resolvió omitir el dato del año 1921 en la tabla 3.4. Esto se debió a que se detectó que en caso de considerar al residual obtenido por el modelo para este año dentro de la muestra, está violaba el supuesto de normalidad de los residuales, es decir, la muestra no respeta la forma de la curva normal, por consiguiente, de considerarlo para la estimación de la función transformada, las estimaciones no serían estadísticamente aceptables y los pronósticos realizados sufrirían de un considerable sesgo debido a los problemas de especificación.

Una vez que la función transformada de la población fue obtenida se sustituyó en el modelo logístico para determinar el modelo demográfico para la población de estudio. Posteriormente se compararon los resultados teóricos obtenidos por el modelo contra los resultados registrados por el Instituto Nacional de Estadística y Geografía, para ello se utilizó un gráfico de dispersión, en el cual se observa la curva logística poblacional, además de introducir el concepto de error relativo, para cuantificar la magnitud y dispersión de los residuales poblacionales obtenido en cada punto temporal de la muestra, dicho concepto resultó tener una medición de los errores relativamente pequeña, por lo que se demuestra que el modelo representa adecuadamente al fenómeno poblacional sustentado en la teoría de la probabilidad y la estadística.

Bajo cada supuesto validado en la metodología de MCO y demostrando el comportamiento asintótico logístico de la población estatal, se demuestra que la relación matemática que representa el comportamiento y pronóstico puntual de la población en el estado poblano está dada por:

$$P_t = 0.9663151 + \frac{6.949239}{1 + e^{6.949239*(-0.0072594t+0.7196245)}}. \qquad (4.1)$$

Donde:

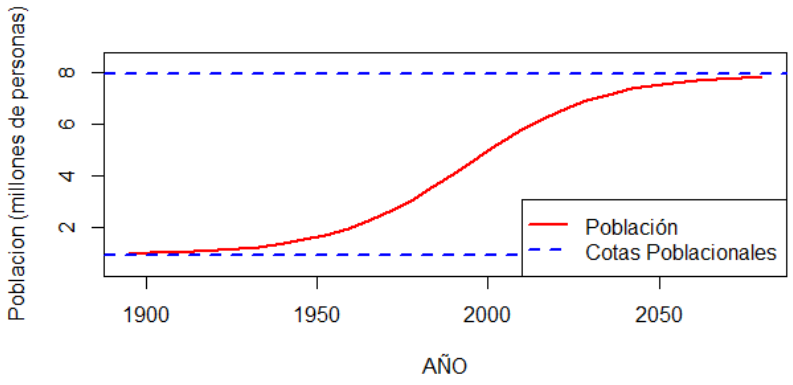
- P_t, denota la media poblacional en el periodo de tiempo t.
- La constante 0.966315 es el valor mínimo de la población del estado dado en unidad de millones.
- La constante 7.915555 denota el máximo población estatal dado en unidades de millones.
- Las contantes -0.00725 y 0.7196245 son los parámetros de rapidez.

Se concluye que el mínimo poblacional del estado de Puebla es de 966,315 habitantes, mientras que la capacidad máxima para la media de la población está dada por 7, 915,555 personas y el modelo matemático para simular el crecimiento poblacional está dado por la ecuación 4.1.

Gracias al software estadístico R se obtiene la figura 4.1, la cual simula el comportamiento demográfico dado por la relación 4.1 en el lapso que consta de 1890 a 2090 donde se aprecia de manera concisa la conjetura de la existencia de un modelo logístico. Se observa que a partir del año 1990 la tasa de crecimiento del modelo alcanza su valor más alto, este concepto tomará un comportamiento descendente, lo que provoca un ligero descenso de la población poblana recibida en cada año posterior.

Se observa además un notable decrecimiento de la tasa de crecimiento alrededor del año 2010 el cuál coincide en términos históricos con el informe del Consejo Nacional de Población que determina una disminución de la tasa global de fecundidad del país. Aspectos como la escolaridad, apertura a la educación sexual en los diferentes niveles educativos del país y los avances en igualdad de derechos a la mujer son algunas variables sociales que bien pueden considerarse importantes para explicar el fenómeno demográfico.

Figura 4.1: *Gráfica de los pronósticos para la demografía estatal ubicando las cotas de población para los años 1890-2090.*



Fuente: Resultados de la ecuación 4.1.

Como cientos de investigaciones lo avalan, la demografía se esta convirtiendo en un tema de sensibilidad extrema. La errónea gestión y distribución demográfica es considerado cada vez más un problema contemporáneo de gran relevancia. (Lomborg, 2014).
Estamos tratando con el que posiblemente sea el detonante principal de una serie de eventos de reacción en cadena que terminará en un escenario nada favorable para cualquier individuo sobre el planeta.Diversos grupos de investigación (universidades, ecologistas, investigadores ambientales, climatológos, etc.), giran ya en torno a sus propias teorías alrededor de las causas y consecuencias de la sobrepoblación. (Tuirán, 1988)

La concientización de la sociedad, la prevención por parte de las organizaciones gubernamentales y la intervención de la ciencia son fundamentales para la preservación de los ecosistemas y la ralentización de los efectos del cambio climatológico causando en manos de la sobrepoblación. Tan solo temas como la sobreproducción agrícola y ganadera son de actual controversia y tema tocado en debates con seriedad mundial.(FAO, 2006)
Eventualmente, las instituciones políticas y legales desempeñaran un papel determinante en la creación de un entorno que pueda alimentar la prosperidad y el crecimiento económico de cada identidad gubernamental del planeta. Por una parte, dejando a un lado las ya marcadas carencias económicas y la desigualdad de nuestro país, es indispensable fortalecer las capacidades del sistema educativo, fomentar la diversidad de ideas en pro de una correcta percepción de la realidad que afronta el planeta. (Parsons, 1991)

La introducción de nuevas visiones en materia de ecología y desarrollo sustentable, así como de nuevas iniciativas y proyectos en educación reproductiva son algunos de los temas que se pondrán en la mesa para su valoración ante los eventos próximos.

Por otro lado, es necesario que las autoridades y organizaciones competentes planifiquen de manera oportuna un plan de contingencia para el desarrollo de los servicios sociales y económicos del país. Los cambios poblacionales por rango de edad son de suma importancia en materia educativa, social y de salud. Los fondos de seguros, pensiones y de indemnizaciones por parte de organizaciones públicas (ISSSTE,IMSS) y privadas están correlacionadas de manera directa con la cantidad de personas que se encuentran en edad productiva y en edad de jubilación (Villa-D, 2019), por lo tanto, su sistema financiero debe ser capaz de soportar la carga económica y administrativa que se requerirá en los próximos años debido a la alteración de la pirámide poblacional ocasionada en mayor medida por la extensión de la esperanza de vida y la reducción de la tasa global de fecundidad. Por esto, el gobierno y las empresas deben tener entre sus prioridades asimilar e interpretar el crecimiento poblacional, así como vigilar la transición dentro y fuera del país, la actividad económica nacional e internacional y las leyes que se formularán alrededor de la demografía.
La expansión poblacional supone un nivel de urbanización mayor, y de la mano,

una disminución de la ecología. Actualmente, la cuidad de Puebla se encuentra en un pulsante proceso de colapso ambiental a raíz de la expansión, por lo que, las decisiones que se tomen tanto en la sociedad como en los gobiernos deben estar fundamentadas en el propósito de obtener los mejores escenarios futuros.

Los efectos de un tema tan generalizado como lo es el descontrol demográfico son tan grandes y tan amplios que cualquier área de la sociedad se verá afectada. Las investigaciones en materia actuarial y demográfica no serán suficientes para estudiar y predecir todas las variables que comenzarán a oscilar en los escenarios adversos que pueden provocarse debido a la mala distribución de la pirámide poblacional mundial. Áreas como la genética, ingenierías en materiales y energías renovables, climatología, restauración ambiental, economía y sociedad, salud, psicología, farmacología y por supuesto, la búsqueda de nuevas alternativas para el cultivo de alimentos y la preservación del agua son investigaciones que bien pueden enlazarse con las actuales investigaciones que buscan predecir el comportamiento humano ante su crecimiento y preservación.

Sin duda, el fenómeno de la sobrepoblación es un tema que pondrá en situaciones adversas a muchos sectores. Todas las ramas de la ciencia deberán moverse con cautela a partir de ahora. Cada paso será de significante importancia en la búsqueda de nuevas maneras de adaptación y sobrevivencia en los años por venir.

Apéndice A

Solución de la ecuación diferencial para el mínimo y máximo de la población.

Bajo la metodología de ecuaciones diferenciales de variables separables se obtiene que:

$$\int \frac{1}{(P-K)(P-K-C)}\,\mathrm{d}P = \int j(t)\,\mathrm{d}t = \lambda(t).$$

Con $P, K, C\ \epsilon\ R^{+}$.

Donde:

- P, es la variable aleatoria que denota el tamaño de la población.
- K, es el mínimo de la población estudiada.
- C, es la diferencia obtenida entre el máximo y mínimo de la población.

Se busca obtener una solución al lado izquierdo de la relación, para esto se utilizará el método de integración por fracciones parciales (Del-Pino, 2019). Con este fin, se recurrirá al artificio algebraico como herramienta para que la expresión:

$$\frac{1}{(P-K)(P-K-C)},$$

pueda formularse como una sustracción de fracciones.

Dada la siguiente relación:

$$\frac{1}{(P-K)(P-K-C)} = \frac{1}{U(U-C).}$$

Donde $U = P - K$.

Por consistencia de la igualdad y propiedad del neutro multiplicativo:

$$\frac{1}{U(U-C)} = \frac{1}{U(U-C)} * 1 = \frac{1}{U(U-C)} * \frac{C^2}{C^2} = \frac{C^2}{U(U-C)C^2.}$$

Por propiedad del neutro aditivo se mantiene la relación:

$$\frac{C^2}{U(U-C)C^2} = \frac{C^2 + (CU - CU)}{U(U-C)C^2} = \frac{CU - CU + C^2}{U(U-C)C^2}$$

Se factoriza al numerador de la expresión racional anterior:

$$\frac{CU - CU + C^2}{U(U-C)C^2} = \frac{CU - C(U-C)}{U(U-C)C^2}.$$

Esta última expresión puede re-escribirse como:

$$\frac{CU - C(U-C)}{U(U-C)C^2} = \frac{CU}{U(U-C)C^2} - \frac{C(U-C)}{U(U-C)C^2}.$$

Simplificando ambas fracciones se obtiene:

$$\frac{CU}{U(U-C)C^2} - \frac{C(U-C)}{U(U-C)C^2} = \frac{1}{C(U-C)} - \frac{1}{CU}.$$

Entonces se afirma que,

$$\int \frac{1}{(P-K)(P-K-C)}\,\mathrm{d}P = \int \left(\frac{1}{C(U-C)} - \frac{1}{CU}\right)\,\mathrm{d}U.$$

Esto último se debe a que se recurre al teorema de cambio de variable (Colegio-de-México, 2012). Una vez postulada la igualdad se busca resolver la integral indefinida.

$$\int \left(\frac{1}{C(U-C)} - \frac{1}{CU}\right)\,\mathrm{d}U = \int \frac{1}{C(U-C)}\,\mathrm{d}U - \frac{1}{CU}\,\mathrm{d}U.$$

Debido a la linealidad de la integral con respecto a los términos constantes, entonces:

$$\int \frac{1}{C(U-C)}\,\mathrm{d}U - \frac{1}{CU}\,\mathrm{d}U = \frac{1}{C}\int \frac{1}{U-C}\,\mathrm{d}U - \frac{1}{C}\int \frac{1}{U}\,\mathrm{d}U.$$

Ingresando funciones primitivas,

$$\frac{1}{C}\int \frac{1}{U-C}\,\mathrm{d}U - \frac{1}{C}\int \frac{1}{U}\,\mathrm{d}U = \frac{1}{C}\ln(U-C) - \frac{1}{C}\ln(U) = \frac{1}{C}\left(\ln(U-C) - \ln(U)\right).$$

Por propiedades de la función logaritmo:

$$\frac{1}{C}\left(\ln(U-C) - \ln(U)\right) = \frac{1}{C}\ln\left(\frac{U-C}{U}\right) = \frac{1}{C}\ln\left(1 - \frac{C}{U}\right).$$

Sabiendo que $U = P - K$, se obtiene que:

$$\int \frac{1}{(P-K)(P-K-C)}\,\mathrm{d}P = \frac{1}{C}\ln\left(1 - \frac{C}{P-K}\right).$$

Al aplicar la función de valor absoluto a los argumentos de funciones logarítmicas con el fin de extender el dominio de la antiderivada se consigue que:

$$\int \frac{1}{(P-K)(P-K-C)}\,\mathrm{d}P = \frac{1}{C}\ln\left(\frac{C}{P-K} - 1\right).$$

Finalmente, se obtiene un resultado para la igualdad expresada por:

$$\lambda(t) = \frac{1}{C}\ln\left(\frac{C}{P-K} - 1\right).$$

Como todos los datos de la muestra no indeterminan la función logaritmo al no resultar en valores negativos, podemos concluir entonces que nuestra solución queda señalada por la relación:

$$\lambda(t) = \frac{1}{C}\ln\left(\frac{C}{P-K} - 1\right).$$

Donde se extiende que:

$$P = K + \frac{C}{1 + e^{C*\lambda(t)}}.$$

Apéndice B

Prueba T

La prueba T, perteneciente al campo de la estadística inferencial, es usada para probar la prueba de hipótesis sobre los coeficientes obtenidos en una regresión lineal simple así como probar la significancia individual de los parámetros en el modelo múltiple (Kyun, 2015) (Montgomery y col., 2012).

Un estadístico basado en la distribución de probabilidad t es usado para probar la hipótesis de que la pendiente e intercepto reales β_0,β_1 son equivalentes a los valores constantes $\hat{\beta}_0$ $\hat{\beta}_1$ respectivamente.

$$H_0 : \beta_0 = \hat{\beta}_0 \quad vs \quad H_a : \beta_0 \neq \hat{\beta}_0.$$

$$H_0 : \beta_1 = \hat{\beta}_1 \quad vs \quad H_a : \beta_1 \neq \hat{\beta}_1.$$

Para el caso del estudio de regresiones lineales múltiples refiriéndose a cualquiera de los k estimadores se afirma que:

$$H_0 : \beta_i = \hat{\beta}_i \quad vs \quad H_a : \beta_i \neq \hat{\beta}_i.[i = 1, 2, \ldots, k.]$$

El estadístico de prueba es expresado por la diferencia entre el parámetro estimado $\hat{\beta}_{(0,1)}$ y el valor a considerar para $\beta_{(0,1)}$ dividida entre la desviación estándar del primero, es decir:

$$T_0 = \frac{\hat{\beta}_{(0,1)} - \beta_{(0,1)}}{se\left(\hat{\beta}_{(0,1)}\right)}.$$

Donde $\hat{\beta}_{(0,1)}$ es el estimador de mínimos cuadrados ordinarios de $\beta_{(0,1)}$ y $se(\hat{\beta}_{(0,1)})$ es su desviación estándar, la cual pueden ser obtenidas por las relaciones:

Para el estimador de la pendiente:

$$se(\hat{\beta}_1) = \sqrt{\frac{\frac{\sum_{i=1}^{n} e_i^2}{n-2}}{\sum_{i=1}^{n}(x_i - \overline{x})^2}}.$$

Para el estimador de la pendiente:

$$se(\hat{\beta}_0) = \sqrt{\frac{\sum_{i=1}^{n} e_i^2}{n-2}\left(\frac{1}{n} + \frac{\overline{x}^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2}\right)}.$$

El estadístico de prueba T_0 sigue una distribución t con $n-2$ grados de libertad, donde n es el número total de observaciones.

$$T_0 \sim t_{(\frac{\alpha}{2},n-2)}.$$

La regla de decisión para la prueba T se presenta como: Se rechaza H_0 si:

$$T_0 \leq -t_{(\frac{\alpha}{2},n-2)} \quad ó \quad T_0 \geq t_{(\frac{\alpha}{2},n-2)}.$$

Es decir, la hipótesis nula H_0 no se rechaza si el valor calculado del estadístico de prueba se encuentra entre los valores:

$$-t_{\frac{(\alpha}{2},n-2)} < T_0 < t_{(\frac{\alpha}{2},n-2)}.$$

Donde $-t_{(\frac{\alpha}{2},n-2)}$ y $t_{(\frac{\alpha}{2},n-2)}$ son los valores críticos para la hipótesis de dos colas. Además $t_{\frac{\alpha}{2},n-2}$ es el percentil de la distribución t correspondiente a la probabilidad acumulada de $\left(1-\left(\frac{\alpha}{2}\right)\right)$ y α es el nivel de significancia.

Si el valor de la constante $\beta_{(0,1)}$ es 0, entonces la prueba de hipótesis es conocida como *prueba de significancia*. No rechazar la prueba $H_0 : \hat{\beta}_1 = 0$ implica que no existe una relación lineal entre la variable independiente x y la variable dependiente y, por lo tanto, se impone un modelo de regresión cuando no existe relación alguna entre ellas.

Apéndice C

Prueba F

También conocida como *Análisis de la varianza (ANOVA)*. Es un método para probar la significancia global de un modelo de regresión (Canavos, 1988),(Montgomery y col., 2012).

Como su nombre lo indica, esta técnica permite utilizar la varianza de los datos observados para determinar si un modelo de regresión puede ser aplicado a la muestra. El análisis de la varianza parte de la siguiente igualdad algebraica:

$$y_i - \overline{y} = (y_i - \hat{y}_i) + (\hat{y}_i - \overline{y}).$$

Donde:

- $y_i - \overline{y}$, es la deviación de cada observación de la variable dependiente respecto a su media. Recibe el nombre de *desviación total*.
- $y_i - \hat{y}_i$, es el error estimado entre el valor observado y el estimado. Recibe el nombre de *desviación debida al error*.
- $\hat{y}_i - \overline{y}$, es la desviación del valor estimado de y con respecto a la media. Depende directamente de la pendiente de la recta obtenida en el modelo de regresión por lo que recibe el nombre de *desviación debida a la regresión*.

Por lo que la partición de la variabilidad queda expresada por:

Desviación Total=Desviación debida al error+Desviación debida a regresión

Puesto que esta igualdad es cierta para todas las observaciones, podemos escribir la expresión dada como:

$$\sum_{i=1}^{n}(y_i - \overline{y})^2 = \sum_{i=1}^{n}((y_i - \hat{y}_i) + (\hat{y}_i - \overline{y}))^2.$$

Al desarrollar la relación anterior, se obtiene que:

$$\sum_{i=1}^{n}(y_i - \overline{y})^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 + \sum_{i=1}^{n}(\hat{y}_i - \overline{y})^2 + 2\sum_{i=1}^{n}(y_i - \hat{y}_i)(\hat{y}_i - \overline{y}).$$

Se observa lo siguiente sobre el último sumando del lado derecho de la expresión:

$$\sum_{i=1}^{n}(y_i-\hat{y}_i)(\hat{y}_i-\overline{y})=\sum_{i=1}^{n}(e_i)(\hat{y}_i-\overline{y})=\sum_{i=1}^{n}(e_i)(\hat{y}_i)-\overline{y}\sum_{i=1}^{n}(e_i)=0.$$

Por supuestos 7 y 3 respectivamente de la metodología de regresión lineal ambos sumandos son nulos. Por lo que:

$$\sum_{i=1}^{n}(y_i-\overline{y})^2=\sum_{i=1}^{n}(y_i-\hat{y}_i)^2+\sum_{i=1}^{n}(\hat{y}_i-\overline{y})^2.$$

Esto representa que la suma de cuadrados de las deviaciones totales es resultado de la adición de la suma de cuadrados de las desviaciones debidas al error y la suma de cuadrados de las desviaciones debidas a la regresión.

$$S.C.TOTAL{=}S.C.ERROR{+}S.C.REGRESIÓN.$$

Recordemos que la varianza muestral puede ser obtenida por:

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

Veamos que el numerador de la varianza muestral es igual a la suma de cuadrados totales. El denominador está asociado al número de grados de libertad asociados a la varianza muestral. Entonces se define al *cuadrado medio total (CMT)* como:

$$CMT=\frac{S.C.TOTAL}{n-1}.$$

De la misma manera se define que:

$$S.C.ERROR=\sum_{i=1}^{n}(y_i-\hat{y}_i)^2=\sum_{i=1}^{n}(e_i)^2.$$

En la bibliografía (Canavos, 1988) (Montgomery y col., 2012) se pueden encontrar una estimación para la suma de cuadrados de los errores la cual se expresa como:

$$S.C.ERROR=\sum_{i=1}^{n}(y_i-\overline{y})^2-\hat{\beta}_1\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y}).$$

Donde x_i,y_i son los valores obtenidos en la muestra y $\hat{\beta}_1$ es el estimador para la pendiente de la regresión lineal simple obtenida bajo la metodología de MCO. El número de grados de libertad asociados con esta expresión es $n-2$.

Entonces el *cuadrado medio de los errores (CME)*, está dado por:

$$CME=\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{n-2}=\frac{\sum_{i=1}^{n}(y_i-\overline{y})^2-\hat{\beta}_1\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{n-2}.$$

Cabe decir que el valor del *cuadrado medio de los errores* es un estimador de la varianza muestral de los errores obtenidos, es decir:

$$S_e^2 = CME = \frac{S.C.ERROR}{n-2}.$$

Así mismo, se define la *suma de cuadrado de la regresión* como:

$$S.C.REGRESIÓN = \sum_{i=1}^{n}(\hat{y}_i - \overline{y})^2.$$

Una vez más, puede encontrarse estimaciones para este concepto que resultan como:

$$S.C.REGRESIÓN = \hat{\beta}_1 \sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y}) = \hat{\beta}_1^2 \sum_{i=1}^{n}(x_i - \overline{x})^2.$$

Los grados de libertad asociados a esta expresión es 1.Por lo tanto, el *cuadrado medio de la regresión (CMR)* está dado por:

$$CMR = \frac{S.C.REGRESIÓN}{1}.$$

Se define el siguiente juego de hipótesis para la Prueba F.

$$H_0 : \beta_1 = 0 \ [vs] \ H_a : \beta_1 \neq 0.$$

El estadístico de prueba está dado por:

$$F_0 = \frac{CMR}{CME} = \frac{\frac{S.C.REGRESIÓN}{1}}{\frac{S.C.ERROR}{n-2}} = \frac{S.C.REGRESIÓN}{S_e^2}.$$

El cual sigue una distribución F con un grado de libertad en el numerador y $n-2$ en el denominador.

$$F_0 \sim F^1_{(n-2,\alpha)}.$$

La regla de decisión para la prueba F está dada por:

$$Rechazar \quad H_0 \quad si \quad F_0 > F_{(n-2,\alpha)}.$$

Donde $F^1_{(n-2,\alpha)}$ es el percentil de la distribución F correspondiente a la probabilidad acumulada de $(1-\alpha)$ y α es el nivel de significancia.

La tabla C.1 representa de manera resumida la formulación de las sumas de cuadrados, los cuadrados medios y el estadístico F del análisis de varianza, así como sus correspondientes grados de libertad.

Para el caso de regresiones lineales múltiples, los conceptos y premisas se basan en los mismos fundamentos del caso de regresión simple, solo se intercambiar la suma de cuadrados de las diferentes desviaciones por sus respectivos valores matriciales y se ajustan los grados de libertad de las fuentes de variación para una prueba F conjunta. La tabla C.2 resume dichos conceptos de la siguiente manera

Tabla C.1: *Análisis de Varianza para el modelo de regresión simple.*

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F_0
Regresión	1	$\hat{\beta}_1 S_{xy}$	$\frac{\hat{\beta}_1 S_{xy}}{1}$	$\frac{\hat{\beta}_1 S_{xy}}{S_e^2}$
Error	$n-2$	$\sum_{i=1}^{n}(y_i-\overline{y})^2 - \hat{\beta}_1 S_{xy}$	S_e^2	
Total	$n-1$	$\sum_{i=1}^{n}(y_i-\overline{y})^2$		

Fuente: Introduction to Linear Regression (Montgomery y col., 2012).

Tabla C.2: *Análisis de Varianza para el modelo de regresión múltiple.*

Fuente de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	F_0
Regresión	k	$\hat{\beta}X'y - \frac{(\sum_{i=1}^{n} y_i)^2}{n}$	$\frac{\hat{\beta}X'y - \frac{(\sum_{i=1}^{n} y_i)^2}{n}}{k}$	$\frac{\frac{S.C.REGRESIÓN}{k}}{\frac{S.C.ERROR}{n-k-1}}$
Error	$n-k-1$	$y'y - \hat{\beta}X'y$	$\frac{y'y-\hat{\beta}X'y}{n-k-1}$	
Total	$n-1$	$y'y - \frac{(\sum_{i=1}^{n} y_i)^2}{n}$		

Fuente: Introduction to Linear Regression (Montgomery y col., 2012).

Apéndice D

Coeficiente de determinación

El coeficiente de determinación (R^2), es comúnmente utilizado como una medida de bondad de ajuste para un modelo de regresión, es decir, mide que tan aproximada es la explicación de la variable independiente x sobre la variable dependiente y en un modelo de regresión lineal simple sobre una muestra dada (Canavos, 1988),(Carter y col., 2011),(Wooldridge, 2009).
Este concepto está altamente enlazado con el coeficiente de correlación ya que mide la asociación entre dos variables, misma que las técnicas de regresión lineal simple tratan de cuantificar.
Basta con usar artificio algebraico para conseguir que:

$$\rho_{xy} = \frac{S_{xy}}{\sqrt{S_x^2}\sqrt{S_y^2}} = \frac{\sqrt{S_x^2}}{\sqrt{S_y^2}} * \frac{S_{xy}}{S_x^2} = \frac{S_x}{S_y}\hat{\beta}_1.$$

Si elevamos al cuadrado a ambos lados de la relación:

$$\rho_{xy}^2 = \left(\frac{S_x}{S_y}\hat{\beta}_1\right)^2 = \frac{\hat{\beta}_1 S_{xy}}{S_y^2} = \frac{\hat{\beta}_1 \sum_{i=1}^n (x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^n (y_i - \overline{y})^2}.$$

Por análisis de la varianza (Apéndice C), tenemos que el numerador de la expresión anterior no es más que la *suma de cuadrados de la regresión* mientras que el denominador es la *suma de cuadrados total*. Al resultado de dicha división es conocido como el coeficiente de determinación.

$$R^2 = \frac{S.C.REGRESIÓN}{S.C.TOTAL}.$$

Usando la Tabla C.1, encontramos una expresión equivalente:

$$R^2 = \frac{S.C.REGRESIÓN}{S.C.TOTAL} = \frac{S.C.TOTAL - S.C.ERROR}{S.C.TOTAL} = 1 - \frac{S.C.ERROR}{S.C.TOTAL}.$$

El estimador R^2 provee la proporción de variabilidad explicada por el modelo en comparación con la variabilidad total. En términos coloquiales, este valor nos explica la *proporción de los datos que están siendo explicados por el modelo de regresión.*

El coeficiente de determinación siempre toma valores entre 0 y 1, debido a que la *S.C. REGRESIÓN* y *S.C. ERROR* no pueden ser mayores que la variabilidad total.

En términos generales, para interpretar al coeficiente de determinación es común multiplicar este resultado por 100, esto nos dará el porcentaje de los datos que son explicados.

Observemos que existe una relación directa entre el coeficiente de correlación y el coeficiente de determinación. Dicha relación se expresa como:

$$\rho_{xy}^2 = R^2.$$

Es decir, el estadístico R^2 es igual al cuadrado del coeficiente de correlación.

Dicho esto, se hacen las siguientes conjeturas:

1. Si el coeficiente de determinación toma valores próximos a 1 existirá mayor evidencia de tener un modelo que explica significativamente a la variable y basándose en la variación de x.

2. Si el coeficiente toma exactamente el valor 1 implicaría que *S.S.ERROR* es nulo y por lo tanto todo punto cae en la muestra caen exactamente sobre la recta estimada.

3. Si el coeficiente toma valores próximos a 0, no hay razón para creer que exista una relación de tipo lineal entre x y y. Esto conllevaría a que *S.C. REGRESIÓN* tomaría un valor muy pequeño en comparación de *S.C. TOTAL*, mientras que *S.C. ERROR* crecería.

4. Se debe tener precaución al momento de interpretar el coeficiente de determinación. El valor de R^2 incrementa con cada término que es añadido a la regresión, aunque estos no contribuyan significativamente al modelo. Así que, un incremento en el valor de R^2 no debe ser considerado como señal para concluir que un modelo es mejor que otro.

En los modelos lineales múltiples, el coeficiente de determinación tiende a ser mayor debido al mayor número de variables explicativas incluida en los modelos. Cada predictor va a explicar una parte de la variabilidad observada en la variable dependiente, por esta razón el coeficiente de determinación no puede utilizarse para comparar modelos con distinto número de predictores.

El coeficiente de determinación ajustado ($R^2_{ajustado}$) introduce una penalización al valor de R^2 por cada predictor que se introduce en el modelo. El coeficiente ajustado permite encontrar el mejor modelo con el menor número de predictores.

El coeficiente de determinación ajustado se contempla como:

$$R^2_{ajustado} = 1 - \left(\frac{S.C.ERROR}{S.C.TOTAL}\right)\left(\frac{n-1}{n-k-1}\right) = R^2 - (1-R^2)\left(\frac{k}{n-k-1}\right).$$

Siendo *S.C. ERROR* la suma de cuadrados de los errores y *S.C. TOTAL* la suma de cuadrados totales (Apéndice C), n el tamaño de muestra y k el número de variables independientes introducidas en el modelo de regresión múltiple.

Apéndice E

Valor P

Se define como el nivel de significancia marginal dispuesta dentro de una prueba de hipótesis representando la probabilidad de que un estadístico calculado ocurra dada una hipótesis nula cierta (Beers, 2020).

Si el valor P cumple con la condición de ser menor que un nivel de significancia dado arbitrariamente, se considera como un resultado estadísticamente significativo. Un valor P menor significa que hay evidencia suficiente a favor de la hipótesis alternativa. Teóricamente, el valor P se define como:

$$\textit{Valor P=P(Valor observado del estadistico ocurre} \quad |H_0 \quad \textit{es cierta).}$$

Al ser una probabilidad, el valor P oscila entre 0 y 1. Se suele decir que niveles altos de para el valor P no permiten rechazar H_0, mientras que niveles bajo permiten rechazarla.

En una prueba de hipótesis, se rechaza la hipótesis nula si el *valor P* asociado al resultado observado es igual o menor que un nivel de significancia α establecido con anticipación.

Una segunda definición para este concepto, la define como la probabilidad, bajo la hipótesis nula H_0, sobre una distribución desconocida F de la variable aleatoria X, para que la variable se observe como un valor igual o mayor que el valor observado.

Si x es el valor observado, dependiendo la manera de interpretarse, un valor mayor o igual al que fue observado puede ser $\{X \geq x\}$ (evento de cola derecha), $\{X \leq x\}$ (evento de cola izquierda) o el evento dando la más pequeña probabilidad entre $\{X \leq x\}$ y $\{X \geq x\}$ (evento de dos colas). Entonces, el Valor P está dado por:

- $P(X \geq x|H_0)$, para un evento de cola derecha.
- $P(X \leq x|H_0)$, para un evento de cola izquierda.
- $2 * \min P(X \geq x|H_0), P(X \leq x|H_0)$, para un evento de dos caras.

Cuanto menor es el valor P, mayor es la importancia esto implica que la hipótesis considerada como cierta no explica adecuadamente la observación H_0 es rechazada

si alguna de estas probabilidades es menor o igual que un valor con umbral relativamente pequeño y fijo, pero arbitrariamente predefinido α llamado el nivel de significancia. Este último, no se deriva de ningún dato de observación y no depende de la hipótesis subyacente, los investigadores comúnmente utilizan niveles de significancia de 0.05,0.01,0.005,0.001.

Apéndice F

Prueba Shapiro-Wilks

Considerada como una prueba para probar la normalidad de una población (Dietrichson, 2019).

La prueba de Shapiro- Wilk (1965) contrasta la hipótesis nula que la muestra $(x_1, x_2, \ldots, x_n)$ se rige bajo una distribución normal utilizando la forma de los polígonos de frecuencias.

El juego de hipótesis de la prueba Shapiro- Wilk se enuncia como:

$$H_0: \quad \textit{La muestra sigue una distribució normal vs } H_a: \quad \textit{No } H_0.$$

$$H_0: x \sim \mathrm{N}(0, \sigma^2) \quad \textit{vs } H_a: \quad \textit{No } H_0.$$

El estadístico de prueba está dado:

$$W = \frac{(\sum_{i=1}^{n} a_i x_{(i)})^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2}$$

Donde:

- $x_{(i)}$ es el i-ésimo estadístico de orden.
- $\overline{x}$ es la media muestral.
- Los coeficientes a_i están dados por:

$$(a_1, a_2, \ldots, a_n) = \frac{m^T V^{-1}}{C}.$$

- Donde C es un vector normal:

$$C = (m^T V^{-1} V^{-1} m)^{\frac{1}{2}}.$$

- El vector m está hecho de los valores esperados de los estadísticos de orden de las variables muestrales aleatorias independientes e idénticamente distribuidas normal estándar.

$$m = (m_1, \ldots, m_n)^T.$$

- V es la matriz de covarianza de los estadísticos de orden normales.

Una forma alternativa de obtener el estadístico de prueba de la Prueba Shapiro-Wilk (1965), se basa en calcular la media y la varianza muestral (S^2) de los datos y las observaciones se ordenan de menor a mayor. A continuación, se calculan las diferencias entre: el primero y el último, el segundo y el penúltimo, el tercero y antepenúltimo, etc. Y se corrigen con unos coeficientes tabulados por los autores Samuel S. Shapiro y Martin B. Wilks.

El estadístico de prueba, bajo esta manera alternativa, se expresa como:

$$W = \frac{D^2}{nS^2}.$$

Donde D es la suma de las diferencias corregidas.

Finalmente, la regla de decisión se expresa como:

Se rechaza H_0 si:

$$W < \theta$$

O bien,

$$Valor\ P \quad < \alpha.$$

Donde θ representa el valor critico proporcionado por la tabla elaborada por Shapiro y Wilks para el tamaño muestral y un nivel de significancia dado α (Barrios y col., 2016).

Apéndice G

Prueba White

Definida como una prueba estadística para demostrar la existencia de heterocedasticidad en una muestra (Carter y col., 2011), (Gujarati & Porter-D., 2010).

Para la presente prueba se hace uso del término función varianza.

Se supone que el modelo de regresión lineal múltiple para el cual se pretende demostrar la existencia de heterocedasticidad es de la forma:

$$E(y_i) = \beta_1 + \beta_2 x_{i2} + \cdots + \beta_k x_{ik}.$$

Se propone un modelo auxiliar que se relaciona con la varianza. Este nuevo modelo está conformado por las variables exploratorias $z_{i2}, z_{i3}, \ldots, z_{is}$ que son todas las posibles formas de $x_{i2}, x_{i3}, \ldots, x_{ik}$. Una forma general de la función varianza es:

$$var(y_i) = h(\alpha_0 + \alpha_1 z_{i1} + \cdots + a_s z_{is}).$$

Donde la varianza cambia por cada observación dependiendo de los valores que tomen las variables exógenas $z's$.

La función varianza solo es relevante cuando la heterocedasticidad es una posibilidad.

Se observa que si $\alpha_1 = \alpha_2 = \cdots = \alpha_s = 0$, entonces:

$$h(\alpha_0 + \alpha_1 z_{i1} + \cdots + a_s z_{is}) = h(\alpha_0).$$

Donde, este último término es una constante.

Lo anterior implica que la varianza no depende de ninguna variable explicativa, en otras palabras, cuando $\alpha_2 = \alpha_3 = \cdots = \alpha_s = 0$ la heterocedasticidad no existe. La varianza es constante.

Entonces se define el juego de hipótesis para la prueba de heterocedasticidad:

$$H_0 : \alpha_2 = \alpha_3 = \cdots = \alpha_s = 0 \quad vs \quad H_a : \quad Alguna\ de\ las \quad a_s \quad es\ diferente\ de\ 0.$$

En términos generales:

$$H_0 : \quad La\ muestra\ es\ homocédastica\ vs \quad H_a : \quad La\ muestra\ es\ heterocédastica.$$

El siguiente componente de la prueba estadística es obtener un estadístico de prueba, así como su distribución.

Como se menciona en la prueba F (Apéndice C) la varianza de un modelo está relacionada con los residuales del mismo. Desde que el coeficiente de determinación R^2 (Apéndice D) se utiliza como una medida de bondad de ajuste, ya que mide la variación de los residuales explicada por las variables auxiliares $z's$, es un buen candidato para la presente prueba.

El econometrista Halbert White (1980) demostró que bajo H_0, el coeficiente de determinación multiplicado por el tamaño de la muestra sigue una distribución Chi-Cuadrada con $S-1$ grados de libertad (se le resta el estimador del intercepto debido a que no es relevante).

Es decir:

$$WHT = N * R^2 \sim \mathcal{X}^2_{(S-1)}.$$

Donde S es el número de variables exógenas $z's$ presentes en el modelo auxiliar y R^2 es el obtenido en la regresión auxiliar.

Halbert White (1980) sugiere que las variables explicadoras auxiliares z's sean las mismas variables $x's$ del modelo de regresión lineal original junto con los cuadrados de cada una de estas, así como todos los posibles productos cruzados entre las variables $x's$.

Además, la variable dependiente del modelo auxiliar no es más que los residuales obtenidos del modelo de regresión lineal original al cuadrado.

Por ejemplo:

Si la forma del modelo de regresión lineal múltiple del cual se busca demostrar la existencia de heterocedasticidad está dada por 3 variables explicativas:

$$E(y_i) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3.$$

Entonces el modelo auxiliar estaría expresado como:

$$\hat{e}^2 = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \alpha_4 x_1 x_2 + \alpha_5 x_1 x_3 + \alpha_6 x_2 x_3 + \alpha_7 x_1^2 + \alpha_8 x_2^2 + \alpha_9 x_3^2.$$

Los residuales estimados al cuadrado de la regresión original han sido considerados como la variable dependiente mientras que a las variables independientes han sido reemplazadas como la prueba de White lo sustenta. En este caso, hay 10 términos como variables independientes, por lo tanto, los grados de libertad toman el valor: $S - 1 = 10 - 1 = 9$. Si la regresión original contiene términos cuadráticos (por ejemplo $x_2 = x_1^2$), algunos términos serán redundantes y deben ser omitidos.

Con un nivel de significancia del 5 %, la regla de decisión se expresa de la siguiente manera:

$$\textit{Se rechaza} \quad H_0 \quad \textit{si:} \quad WHT \geq \mathcal{X}^2_{(0.95,S-1)} \quad \textit{ó Valor } P \quad < \alpha.$$

Por lo tanto, si se concluye el rechazo de la hipótesis nula, se está comprobando entonces la existencia de heterocedasticidad.

Apéndice H

Prueba Durbin-Watson

Prueba utilizada para la detección de la existencia de autocorrelación entre los datos de una muestra (Kenton, 2019).

Respecto al estudio de regresiones lineales, decimos que existe autocorrelación cuando el término de error de un modelo está correlacionado consigo mismo a través del tiempo, tal que $\text{cov}(e_i, e_j) \neq 0$. Comprobar este concepto es de suma importancia en el estudio de regresiones lineales ya que viola un supuesto indispensable de la metodología.

La existencia de autocorrelación en los residuos es fácilmente identificable obteniendo las funciones de autocorrelación (ACF) y autocorrelación parcial (ACP) de los errores mínimo cuadráticos obtenidos en la estimación. Si dichas funciones corresponden a una distribución normal con media 0 y varianza constante, se constatará la ausencia de correlación entre los residuos. Sin embargo, se pueden utilizar diversos contrastes y pruebas para la autocorrelación, siendo la más utilizada la prueba Durbin-Watson (1950).

Se pretende demostrar la ausencia de una relación que explique a un residual dado con su antecesor, es decir, no exista una correspondencia tal que:

$$\hat{e}_t = \rho \hat{e}_{t-1} + u_t.$$

Donde u_t es el término aleatorio de la observación t del nuevo modelo de regresión.

El juego de hipótesis para la prueba Durbin-Watson:

$$H_0 : \rho = 0 \quad vs \quad H_a : \rho \neq 0.$$

O de manera análoga:

$$H_0 : \quad \textit{Hay ausencia de autocorrelación vs} \quad H_a : \quad \textit{Existe autocorrelación.}$$

Dicho contraste se basa en el cálculo del siguiente estadístico de prueba:

$$d = \frac{\sum_{t=2}^{n} (\hat{e}_t - \hat{e}_{t-1})^2}{\sum_{t=1}^{n} \hat{e}_t^2}.$$

Dicho estadístico contiene valores entre 0 y 4, siendo los valores cercanos a 2 indicios de ausencia de autocorrelación de primer orden.

Se ha demostrado que el valor esperado del estadístico cuando no existe la correlación entre los residuales corresponde a la relación:

$$d \approx 2 + \frac{2(k-1)}{n-k}.$$

Para la toma de decisión de esta prueba los estadistas autores de la prueba, Durbin y Watson, derivaron dos distribuciones **dL** y **dU**, las cuales no toman relación con el número de variables independientes y entre estas concentran la distribución verdadera del estadístico d.

Para una prueba de dos colas para autocorrelación ($|\rho| > 0$) y un nivel de significancia dado, se consideran las siguientes reglas de decisión:

1. Si $d < dL_{(\frac{\alpha}{2}),n,k}$ ó $4 - d < dL_{(\frac{\alpha}{2}),n,k}$ se rechaza la hipótesis nula. Existe autocorrelación.

2. Si $d > dL_{(\frac{\alpha}{2}),n,k}$ ó $4 - d > dL_{(\frac{\alpha}{2}),n,k}$ no se rechaza la hipótesis nula. No existe autocorrelación.

3. En otro caso, la prueba no es concluyente.

O bien, Se rechaza H_0 si $\alpha >$ *valor P* .

Apéndice I

Validación de los supuestos de regresión

En este apartado se realizan las pruebas estadísticas para probar los supuestos de regresión bajo la metodología de mínimos cuadrados ordinarios (MCO) de los modelos utilizados en esta obra. Todos los cálculos y gráficas fueron realizados usando el software estadístico R versión 3.6.2 (R-STUDIO, 2019).

Modelo para la aproximación de la parábola obtenida por la relación pendiente - población media.

Supuesto 1: Estructura del modelo.

Con la teoría y metodología antes mencionada aplicada a la muestra se sustenta que:

$$\nabla_t^P = -0.007699 PM_{t_i}^2 + 0.06838 PM_{t_i} - 0.05888 + \pi_i.$$

Supuesto 2: Existencia de variabilidad de la variable independiente.

Utilizando el comando *factor* del software estadístico se obtiene el número de valores diferentes que toma la variable población media. El resultado arrojado para dicho supuesto es expresado por:

$$factor(PM_{t_i}) = 14 - niveles.$$

Es decir, con una muestra de tamaño 14, se obtienen 14 diferentes valores para la variable independiente. Esto ocurre para ambas variables independientes del modelo. Por lo tanto, la variable independiente es conocida y sus valores tienen el supuesto de variabilidad.

Supuesto 3: La suma de los residuales, así como su esperanza matemática es nula.

Se hace uso de los comandos *sum* y *mean* para obtener la suma y promedio de los residuales obtenidos para la regresión respectivamente. La tabla I.1 permite

visualizar los residuales obtenidos para este modelo de regresión considerando los errores aleatorios presentados en la relación 3.6.

Tabla I.1: *Tabulación de los residuales para la regresión de segundo grado.*

Observación (x_i)	Residuo (π_i)
1	-0.0038
2	0.0030
3	-0.0121
4	0.0076
5	0.0012
6	0.0086
7	-0.0044
8	-0.0023
9	0.0086
10	-0.0113
11	0.0067
12	-0.0012
13	-0.0036
14	0.0030

Fuente: Obtenidos por el modelo de regresión 3.6

Al hacer uso del comando *sum* del software estadístico se obtienen el siguiente resultado:

Suma de los residuales ($\hat{\pi}_i$):	-2.168404*10^{-18}

Con respecto a la esperanza matemática de los residuales se considera su media aritmética como estimador. Se hace uso del comando *mean* de R para obtener que:

Media de los residuales ($\hat{\pi}_i$):	-1.548709*10^{-19}

Supuesto 4: La covarianza muestral entre los regresores y los residuales estimados es cero.

Al ser una regresión con términos con segundo grado se hace uso de ambas variables independientes, así como de los residuales mostrados en la tabla I.2.

Haciendo uso de la covarianza muestral, dada por I.1, se busca obtener un resultado muy próximo a 0 para validar este supuesto.

$$\operatorname{cov}(x_i, y_i) = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{n}. \tag{I.1}$$

Se hacen los cálculos correspondientes dados los datos en la tabla I.2 bajo la relación I.1 considerando a los residuales con la primera variable independiente. Se hace uso

Tabla I.2: *Relación entre las variables regresoras y los residuales obtenidos.*

$PM_{t_i}^2$	PM_{t_i}	$\hat{\pi}_i$
1.0469	1.0232	-0.0038
1.1265	1.0614	0.0030
1.1306	1.0633	-0.0121
1.1831	1.0877	0.0076
1.4946	1.2225	0.0012
2.1323	1.4602	0.0086
3.2394	1.7998	-0.0044
5.0222	2.2410	-0.0023
8.5729	2.9280	0.0086
13.9644	3.7369	-0.0113
19.1427	4.3752	0.0067
23.5276	4.8505	-0.0012
31.1529	5.5815	-0.0036
35.6929	5.9744	0.0030

Fuente: Obtenidos por el modelo de regresión 3.6

$\text{cov}(PM_{t_i}^2, \hat{\pi}_i)$:	$1.466279{*}10^{-18}$

del comando *cov* para estimar la covarianza muestral entre estas dos variables.

Se obtiene que:

Respecto a la relación entre los residuales y la segunda variable independiente, se estima la covarianza muestral entre dichas variables.

Una vez más, se hace uso del comando *cov* para estimar la covarianza muestral entre estas variables dado como resultado:

$\text{cov}(PM_{t_i}, \hat{\pi}_i)$:	$2.327386{*}10^{-19}$

Se valida este supuesto.

Supuesto 5: Homocedasticidad.

Para este supuesto se hace uso de la Prueba estadística de White (Apéndice G) donde se demuestra la ausencia de heterocedasticidad para este modelo de regresión.

Se recuerda la prueba de hipótesis para dicha prueba:

H_0 : *La muestra es homocédastica. vs* H_a : *La muestra es heterocédastica.*

Ahora, se calcula el estadístico de prueba para la prueba de White:

$$WHT = N * R^2.$$

Donde R^2 es el coeficiente de determinación del modelo auxiliar dado por:

$$\hat{\pi}_i^2 = J + \delta_1 PM_{t_i}^4 + \delta_2 PM_{t_i}^3 + \delta_4 PM_{t_i}^2 + \delta_5 PM_{t_i}. \tag{I.2}$$

Entonces, el estadístico de White toma el valor:

$$WHT = 14 * 0.30454 = 4.26369$$

Considerando el valor crítico de la distribución *Chi-cuadrada* con 4 grados de libertad a un nivel de significancia del 5 % se obtiene:

$$\mathcal{X}^2_{(.95,4)} = 9.487$$

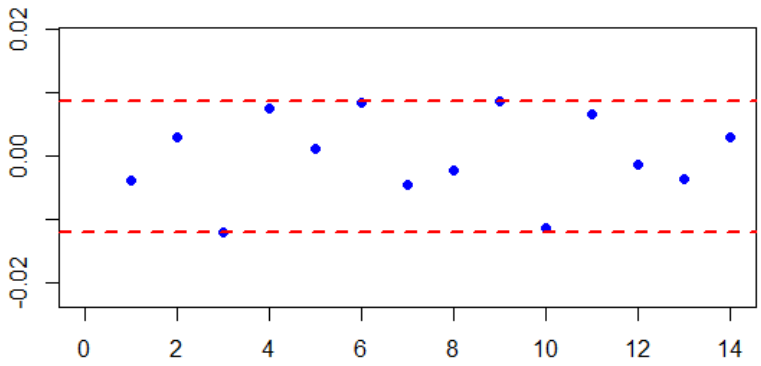
Se obtiene que $WHT < \mathcal{X}^2_{(.95,4)}$, por lo tanto, no se rechaza la hipótesis nula.

Llamando a la prueba White en el programa R se encuentra los siguientes resultados: Se aprecia un *valor P* mayor al nivel de significancia, por lo tanto, existe

WHT	**DF**	**Valor P**
4.2636	4	0.3715

evidencia estadística para probar la ausencia de heterocedasticidad. Los residuales siguen una varianza constante. Como apoyo visual se muestra la gráfica de los residuales obtenidos en la figura I.1.

Figura I.1: *Gráfica de los residuales obtenidos para el modelo de regresión parabólico.*



Obtenidos por el modelo de regresión dado por la ecuación 4.6.

Supuesto 6: La evaluación del punto $\overline{X}$ en el modelo ajustado resulta en el punto $\overline{y}$.

Utilizando el comando *mean* se obtiene las medias de cada variable de este modelo de regresión:

$$\overline{PM^2} = 10.602 \quad \overline{PM_{t_i}} = 2.743257 \quad \overline{\nabla_t^P} = 0.04707343$$

Tomando la ecuación 3.6, se considera que:

$$\hat{\nabla}_M^P = -0.007699(10.602)^2 + 0.06838(2.743257) - 0.05888 = 0.047$$

Con un error de $2.0816 * 10^{-17}$ se demuestra que $\hat{\nabla}_M^P = \overline{\nabla}_t^P$. Este supuesto se valida como cierto.

Supuesto 7: La covarianza muestral entre $\hat{\nabla}_t^P$ y $\hat{\pi}_i$ es nula.

Se busca ahora la covarianza existente entre la variable dependiente y los residuales obtenidos de la regresión parabólica. La tabla I.3 muestra los datos obtenidos de dichas variables.

Tabla I.3: *Tabulación de los valores obtenidos para la variable dependiente y los residuales.*

$\hat{\nabla}_t^P$	$\hat{\pi}_i$
0.0030	-0.0038
0.0050	0.0030
0.0051	-0.0121
0.0064	0.0076
0.0132	0.0012
0.0245	0.0086
0.0392	-0.0044
0.0557	-0.0023
0.0753	0.0086
0.0891	-0.0113
0.0929	0.0067
0.0917	-0.0012
0.0829	-0.0036
0.0748	0.0030

Fuente: Obtenidos por el modelo de regresión 3.6

Se hace uso de la relación I.1 para calcular la covarianza muestral entre estas variables. Usando el comando *cov* se obtiene el resultado:

$\text{cov}(\hat{\nabla}_t^P, \hat{\pi}_i)$**:**	$5.16262{*}10^{-21}$

Por lo tanto, se puede justificar a este supuesto como válido.

Supuesto 8: Ausencia de autocorrelación.

Para probar la validez de este supuesto se hace uso de la prueba Durbin-Watson (Apéndice H). Se intenta demostrar la ausencia de autocorrelación en la muestra de los residuales obtenidos.

Se recuerda el juego de hipótesis para dicha prueba:

H_0 : *Hay ausencia de autocorrelación vs* H_a : *Existe autocorrelación*

Usando la prueba Durbin- Watson en el software estadístico se obtienen los resultados:

dU	**dL**
0.787	1.409

Con $k = 2$ como el número de regresores y un tamaño de muestra $n = 14$, se consideran los valores críticos de la distribución Durbin-Wartson con un valor de significancia de 0.05:

d	**k**	**Valor P**
3.1502	2	0.07186

Se observa que $d > dU$. Además, se cuenta con un valor P mayor a 0.05, entonces, no se rechaza la hipótesis nula. No existe evidencia de autocorrelación entre los residuales.

Supuesto 9: Los residuales siguen una distribución normal con media 0 y varianza constante.

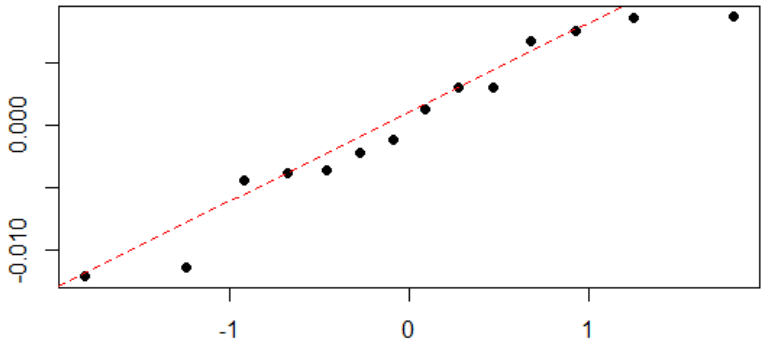
Para este supuesto se llama a la prueba de Shapiro-Wilk (Apéndice F) para conocer si la distribución de los residuales se comporta como una normal centrada en 0. Además se utilizan apoyos gráficos como el histograma superpuesto y el gráfico Q-Q (cuantil-cuantil).

En primer lugar, se realizan los gráficos de normalidad para los residuales estimados $\hat{\pi}_i$.

1. Grafico Q-Q. El gráfico Q-Q (cuantil-cuantil) es un gráfico de probabilidad que compara los valores ordenados de una variable aleatoria arbitraria con los cuantiles de una distribución normal teórica. Se espera que los puntos resultantes se acerquen a una recta diagonal.

   Usando la herramienta estadística R se obtiene la figura I.2, la cual muestra la recta Q-Q de la distribución normal.

Figura I.2: *Gráfico Q-Q de los residuales obtenidos para la regresión de segundo.*



Obtenidos por el modelo de regresión dado por I.2

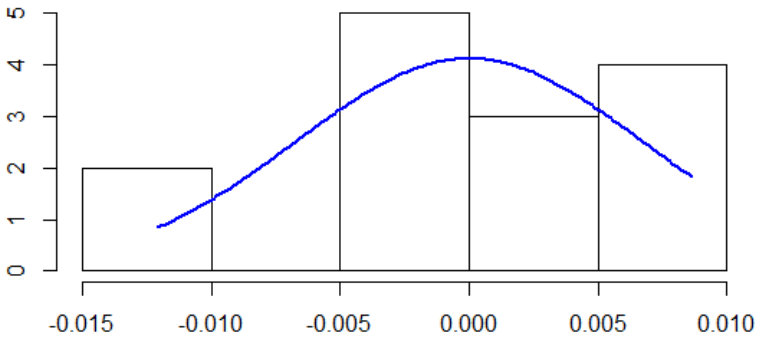
Se observa una tendencia lineal de los cuantiles resultantes de la distribución de los residuales, misma que se aproxima a los cuantiles de una distribución normal.

2. Histograma superpuesto.

   Consiste en representar los datos mediante un histograma y superponer la curva que describe una distribución con la misma media y desviación estándar que la muestra a estudiar.

   Realizando el gráfico en el software se obtiene el histograma resultante. La figura I.3 muestra la curva obtenida.

Figura I.3: *Histograma de los residuales obtenidos para la regresión de segundo grado.*



Obtenidos por el modelo de regresión dado por I.2

Para dar un veredicto formal sobre la normalidad de los residuales se hace uso de la prueba estadística.

3. Prueba Shapiro –Wilk.

   Recordando el juego de hipótesis para esta prueba:

$$H_0: \quad \textit{La muestra sigue una distribución normal vs} \quad H_a: \quad \textit{No } H_0.$$

   El siguiente paso es calcular el estadístico de prueba. Usando R se encuentra los siguientes resultados:

W	Valor P
0.93111	0.3163

   Consultando las tablas de los autores Shapiro y Wilk de valores críticos para la prueba. El valor crítico y el nivel de significancia se dan por:

θ	α
0.024	0.05

   Se evidencia que $W > \theta$, así como que $valorP > \alpha$. Por lo tanto, se concluye que los residuales siguen una distribución normal, con media 0 (se probó en el supuesto 3) y varianza constante (probado en el supuesto 5).

Modelo para la aproximación de la función transformada de la población.

Supuesto 1:Estructura del modelo.

Con la teoría y metodología antes mencionada aplicada a la muestra se sustenta que:

$$\lambda(t) = -.0072594t + 0.7196245 + e_i.$$

Supuesto 2: Existencia de variabilidad de la variable independiente.

Utilizando el comando *factor* del software estadístico se obtiene el número de valores diferentes que toma la variable *periodo*.El resultado arrojado para dicho supuesto es expresado por:

$$factor(t) = 15 \quad niveles \quad .$$

Es decir, con una muestra de tamaño 15, se obtienen 15 diferentes valores para la variable independiente. Esto ocurre para ambas variables independientes del modelo. Por lo tanto, la variable independiente es conocida y sus valores tienen el supuesto de variabilidad.

Supuesto 3: La suma de los residuales, así como su esperanza matemática es nula.

Se hace uso de los comandos *sum* y *mean* para obtener la suma y promedio de los residuales obtenidos para la regresión especificada respectivamente.

La tabla I.4 permite visualizar los residuales obtenidos para este modelo de regresión considerando los errores aleatorios presentados en la relación (Zill, 2011).

Tabla I.4: *Tabulación de los residuales obtenidos por el modelo de la función transformada.*

x_i	$\hat{e}_i$
1	-0.0345
2	0.0124
3	-0.0467
4	0.0531
5	0.0393
6	0.0042
7	0.0076
8	0.0054
9	-0.0088
10	-0.0038
11	-0.0089
12	-0.0106
13	-0.0011
14	-0.0017
15	-0.0056

Fuente: Obtenidos por el modelo de regresión 3.16

Al hacer uso del comando sum del software estadístico se obtienen el siguiente resultado:

Suma de los residuales $\hat{e}_i$:	$1.2143*10^{-17}$

Con respecto a la esperanza matemática de los residuales se considera su media aritmética como estimador. Se hace uno del comando *mean* de R para obtener que:

Media de los residuales $\hat{e}_i$:	$8.0956*10^{-18}$

Supuesto 4: La covarianza muestral entre los regresores y los residuales estimados es cero.

Haciendo uso de la covarianza muestral, dada por I.1, se busca obtener un resultado muy próximo a 0 para validar este supuesto.

Tabla I.5: *Relación entre la variable regresora y los residuales obtenidos del modelo lineal.*

t	$\hat{e}_i$
0	-0.0345
5	0.0124
15	-0.0467
35	0.0531
45	0.0393
55	0.0042
65	0.0076
75	0.0054
85	-0.0088
95	-0.0038
100	-0.0089
105	-0.0106
110	-0.0011
115	-0.0017
120	-0.0056

Fuente: Obtenidos por el modelo de regresión 3.16

Se hace uso del comando *cov* para estimar la covarianza muestral entre estas dos variables. Se obtiene que:

cov(t,$\hat{e}_i$):	-3.69442*10^{-17}

Se valida este supuesto.

Supuesto 5: Homocedasticidad

Para este supuesto se hace uso de la Prueba estadística de White (Apéndice G) donde se demuestra la ausencia de heterocedasticidad para este modelo de regresión.

Se recuerda la prueba de hipótesis para dicha prueba:

$H_0:$ *La muestra es homocédastica. vs* $H_a:$ *La muestra es heterocédastica.*

Ahora, se calcula el estadístico de prueba para la prueba de White:

$$WHT = N * R^2.$$

Donde R^2 es el coeficiente de determinación del modelo auxiliar dado por:

$$\hat{e}_i^2 = \phi_i t^2 + \phi_2 t + p. \tag{I.3}$$

Entonces, el estadístico de White toma el valor:

$$WHT = 15 * 0.39 = 5.85$$

Considerando el valor crítico de la distribución *Chi- cuadrada* con 2 grados de libertad a un nivel de significancia del 5 % se obtiene:

$$\mathcal{X}^2_{(.95,2)} = 5.99$$

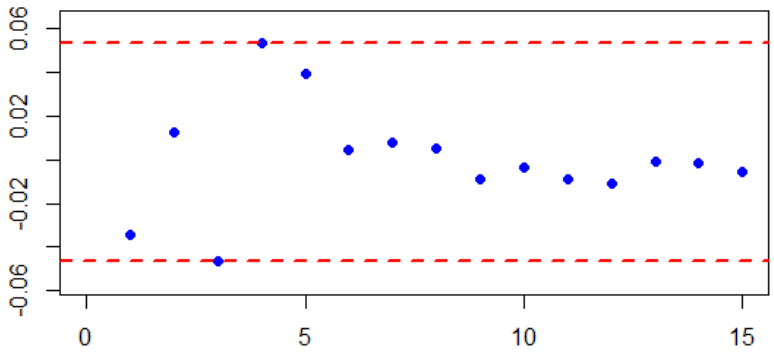
Se obtiene que $WHT < \mathcal{X}^2_{(.95,2)}$, por lo tanto, no se rechaza la hipótesis nula.

Llamando a la prueba White en el programa R se encuentra los siguientes resultados: Se aprecia un valor P mayor al nivel de significancia, por lo tanto, existe

WHT	**DF**	**Valor P**
5.8541	2	0.05356

evidencia estadística para probar la ausencia de heterocedasticidad. Los residuales siguen una varianza constante. Como apoyo visual se muestra la gráfica de los residuales obtenidos en la tabla I.5.

Figura I.4: *Gráfica de los residuales obtenidos para el modelo de la función transformada.*



Obtenidos por el modelo de regresión dado por I.3

Supuesto 6: La evaluación del punto $\overline{x}$ en el modelo ajustado resulta en el punto $\overline{y}$.

Utilizando el comando *mean* se obtiene las medias de cada variable de este modelo de regresión:

$$\bar{t} = 68.3333 \quad \overline{\lambda(t)} = 0.22356$$

Tomando la ecuación 3.18, se considera que:

$$\hat{\lambda(t)}_M = -.0072594(68.3333) + 0.7196245 = 0.2235665$$

Con un error de 0 se demuestra que $\hat{\lambda}(t)_M = \overline{\lambda(t)}$.

Este supuesto se valida como cierto.

Supuesto 7: La covarianza muestral entre $\hat{\lambda(t)}$ y $\hat{e}_i$ es nula.

Se busca ahora la covarianza existente entre la variable dependiente y los residuales obtenidos de la regresión lineal. La tabla I.6 muestra los datos obtenidos de dichas variables.

Tabla I.6: *Relación entre la variable regresora y los residuales obtenidos por la función transformada.*

$\hat{\lambda(t)}$	$\hat{e}_i$
0.7196	-0.0345
0.6833	0.0124
0.6107	-0.0467
0.4655	0.0531
0.3930	0.0393
0.3204	0.0042
0.2478	0.0076
0.1752	0.0054
0.1026	-0.0088
0.0300	-0.0038
-0.0063	-0.0089
-0.0426	-0.0106
-0.0789	-0.0011
-0.1152	-0.0017
-0.1515	-0.0056

Fuente: Obtenidos por el modelo de regresión 3.16

Se hace uso de la relación I.1 para calcular la covarianza muestral entre estas variables.

Usando el comando *cov* se obtiene el resultado:

$\mathbf{cov}(\overset{P}{t}, \hat{\pi}_i)$**:**	-1.864*10^{-19}

Por lo tanto, se puede justificar a este supuesto como válido.

Supuesto 8: Ausencia de autocorrelación.
Para probar la validez de este supuesto se hace uso de la prueba Durbin-Watson (Apéndice H). Se intenta demostrar la ausencia de autocorrelación en la muestra de los residuales obtenidos.

Se recuerda el juego de hipótesis para dicha prueba:

H_0 : *Hay ausencia de autocorrelación vs* H_a : *Existe autocorrelación*

Usando la prueba Durbin- Watson en el software estadístico se obtienen los resultados:

dU	**dL**
0.949	1.222

Con $k = 1$ como el número de regresores y un tamaño de muestra $n = 15$, se consideran los valores críticos de la distribución Durbin-Wartson con un valor de significancia de 0.05:

d	**k**	**Valor P**
2.0988	1	0.9102

Se observa que $d > dU$. Además, se cuenta con un valor P mayor a 0.05, entonces, no se rechaza la hipótesis nula. No existe evidencia de autocorrelación entre los residuales.

Supuesto 9: Los residuales siguen una distribución normal con media 0 y varianza constante.

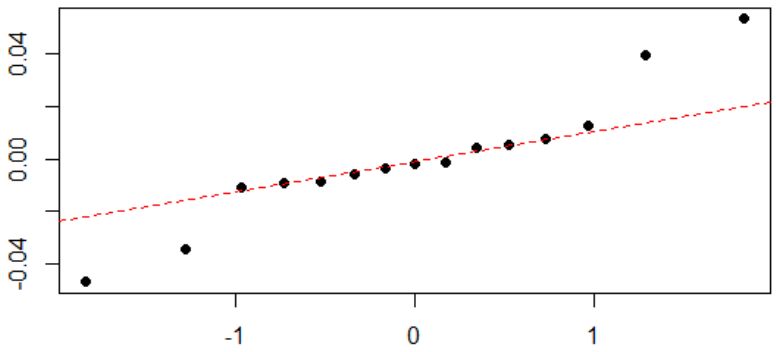
Para este supuesto se llama a la prueba de Shapiro-Wilk (Apéndice F) para conocer si la distribución de los residuales se comporta como una normal centrada en 0. Además se utilizan apoyos gráficos como el histograma superpuesto y el gráfico Q-Q (cuantil-cuantil).

En primer lugar se realizan los gráficos de normalidad para los residuales estimados

1. Gráfico Q-Q. El gráfico Q-Q (cuantil-cuantil) es un gráfico de probabilidad que compara los valores ordenados de una variable aleatoria arbitraria con los cuantiles de una distribución normal teórica. Se espera que los puntos resultantes se acerquen a una recta diagonal.

   Usando la herramienta estadística R se obtiene la figura I.5, la cual muestra la recta Q-Q de la distribución normal.

Figura I.5: *Gráfico Q-Q de los residuales obtenidos para la regresión lineal de la función transformada.*

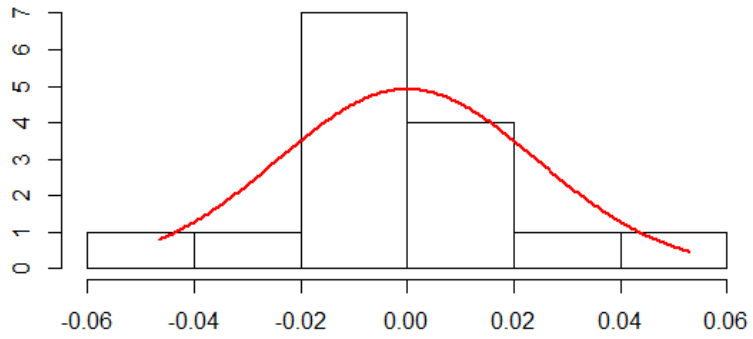


Obtenidos por el modelo de regresión dado por 3.16

2. Histograma superpuesto. Consiste en representar los datos mediante un histograma y superponer la curva que describe una distribución normal con la misma media y desviación estándar que la muestra a estudiar.

   Realizando el gráfico en el software se obtiene el histograma resultante. La figura I.6 muestra la curva obtenida.

Figura I.6: *Histograma de los residuales obtenidos para la regresión lineal..*



Obtenidos por el modelo de regresión dado por 3.16

Para dar un veredicto formal sobre la normalidad de los residuales se hace uso de la prueba estadística.

3. Prueba Shapiro –Wilks.

Recordando el juego de hipótesis para esta prueba:

$$H_0: \quad \textit{La muestra sigue una distribución normal vs} \quad H_a: \quad \textit{No } H_0.$$

El siguiente paso es calcular el estadístico de prueba. Usando R se encuentra los siguientes resultados:

W	**Valor P**
0.92192	0.206

Consultando las tablas de los autores Shapiro y Wilk de valores críticos para la prueba. El valor crítico y el nivel de significancia se dan por:

θ	α
0.01	0.05

Se evidencia que $W > \theta$, así como que $valorP > \alpha$. Por lo tanto, se concluye que los residuales siguen una distribución normal, con media 0 (se probó en el supuesto 3) y varianza constante (probado en el supuesto 5).

Apéndice J

Código en R

Punto de descarga para la base de datos: https://bit.ly/2Q2q2Dw.

```
#1) OBTENCIÓN DEL MODELO POBLACIONAL

# Carga de la base datos año-población.
datos<-read.csv("TESIS.csv")
año<-datos$AÑO
poblacion<-datos$POBLACIÓN


#Grafica para el crecimiento demógrafico:
plot(año,poblacion, type="l", col="black", lty=2,lwd=1, xlim=c(1895,
2020), ylim=c(min(poblacion), max(poblacion)), xlab="AÑO ",
ylab="POBLACIÓN (MILLONES)", main= "POBLACIÓN DEL ESTADO DE PUEBLA")
points(año, poblacion, pch=19, col=12)

#Carga de variables población media- pendiente
poblacion_media<-c(datos$PUNTOS.MEDIOS[1:12],datos$PUNTOS.MEDIOS[14:15])
pendiente<-c(datos$PENDIENTE[1:12], datos$PENDIENTE[14:15])
regresion_1<- lm(pendiente~I(poblacion_media^2)+poblacion_media)

 #Grafica para la curva parabólica población media- pendiente
plot(poblacion_media,pendiente,pch=19, col="red", lwd="1", xlim=c(1, 8),
ylim=c(-0.01,0.1), xlab="Población media (Millones)", ylab="Pendiente",
main= "POBLACIÓN MEDIA  VS PENDIENTES")
x<-seq(1,8,length.out = 50)
r<-regresion_1$coefficients[2]*x^2+regresion_1$coefficients[3]*x+
regresion_1$coefficients[1]
points(x,r,type ="l", lty= 2, col="blue", pch=19, lwd=2)
```

```
#Obtención mínimo y máximo de la población.
#Estimadores obtenidos para la curva parabólica:
A<-regresion_1$coefficients[2]
B<-regresion_1$coefficients[3]
D<-regresion_1$coefficients[1]

#Mínimo y máximo de la población:
minimo<-(-B/(2*A))+(sqrt(B^2-4*A*D)/(2*A))
maximo<-(-B/(2*A))-(sqrt(B^2-4*A*D)/(2*A))
K<-minimo
C<-maximo-minimo

#Estimar valores para la función transformada
transformada1<-(1/C)*log((C/(poblacion-K))-1)

#Modelo para la función desconocida respecto al tiempo.
periodo<-c(datos$PERIODO[1:3],datos$PERIODO[5:16])
transformada<-c(transformada1[1:3], transformada1[5:16])
regresion_2<-lm(transformada~periodo)

#Coeficientes de la regresión
alfa<-regresion_2$coefficients[2]
beta<-regresion_2$coefficients[1]

#Grafica para la función transformada
x<-seq(0,120, length.out = 100)
 r<-regresion_2$coefficients[1]+x*regresion_2$coefficients[2]
plot(x,r, type="l", col="black", lty=2,lwd=1, xlim=c(min(periodo),
max(periodo)), ylim=c(min(transformada), max(transformada)), xlab=
"Tiempo", ylab="Valor transformada", main= "FUNCIÓN TRANSFORMADA DE
LA POBLACIÓN")

  points(periodo, transformada, pch=19, col=12)

#Modelo de Predicción
periodo<-datos$PERIODO
poblacion<-datos$POBLACIÓN
año<-datos$AÑO
modelo<-K+(C/(1+exp(C*(periodo*alfa+beta))))

#Estimación de la medida de los errores ajustados.
diferencia<-poblacion-modelo
(MEA<-sum(diferencia^2))

#Contraste de gráficas entre la muestra y los resultados estimados
```

```
plot(año, modelo, type = "l", ylim = c(1,7), col = "red", lty = 2,
lwd=2,ylab="Poblacion (millones de personas)", xlab="AÑO" )
points(año,poblacion, col="blue", pch=19, lwd=1)
legend("topleft", c("Población Observada", "Población Estimada"),
col = c("blue","red"), text.col = "black", lty = c(-1, 2), pch =
c(19, NA),merge = TRUE)

#RESULTADOS
library(stargazer)
diferencia<-round((poblacion-modelo)*1000000,0)
Poblacion_M<-poblacion*1000000
Modelo_M<-round(modelo*1000000,0)
resultados<-cbind(año,Poblacion_M,Modelo_M,diferencia)
stargazer(resultados, title = "RESULTADOS DEL MODELO", type="text")

# Gráfica de la población estimada con las cotas del máximo y mínimo.
prons<-round(seq(1895,2080, length.out =30 ),0)-1895
a1<-prons+1895
m<-K+(C/(1+exp(C*(prons*alfa+beta))))
plot(a1, m, type = "l", ylim = c(K-.5,maximo+.5), col = "red", lwd=2,
ylab="Poblacion (millones de personas)", xlab="AÑO" )
abline(h=K, lty=2, col="black", lwd=2)
abline(h=maximo, lty=2, col="black", lwd=2)
legend("bottomright", c("Población", "Cotas Poblacionales"), col =
c("red","black"), text.col = "black", lty = c(1, 2), pch = c(NA, NA),
merge = TRUE, lwd=c(2,2))

#Funcion poblacional
poblacion_estatal<-function(año){
evaluar<-año-1895
(resultado<-K+(C/(1+exp(C*(evaluar*alfa+beta)))))}

#Pronósticos para los próximos 10 años
secuencia<-seq(2020,2030,length.out = 11)
pron<-poblacion_estatal(secuencia)
pronosticos<-pron*1000000
t<-as.data.frame(cbind(secuencia, pronosticos))
names(t)<-c("AÑO", "POBLACIÓN ESTIMADA")
t

# Gráfica para los próximos 10 años
prons<-round(seq(1950,2030, length.out =30 ),0)-1895
a1<-prons+1895
m<-K+(C/(1+exp(C*(prons*alfa+beta))))
plot(a1, m, type = "l", ylim = c(K-.5,maximo+.5), col = "red"
```

```
, lwd=2,ylab="Poblacion (millones de personas)", xlab="AÑO" )
abline(h=K, lty=2, col="black", lwd=2)
abline(h=maximo, lty=2, col="black", lwd=2)


#2) VALIDACIÓN DE LOS MODELOS DE REGRESIÓN

#Modelo 1: Modelo para la curva parabolica.

#Supuesto 2: Variabilidad en la variable independiente
          factor(poblacion_media)

# Supuesto 3
 residuales1<- regresion_1$residuals
#La suma de los residuales es 0.
sum(residuales1)
#La media de los residuales es 0.
mean(residuales1)

#Supuesto 4: La covarianza muestral entre los regresores y los
residuales es 0.
poblacionm<-c(datos$PUNTOS.MEDIOS[1:12],datos$PUNTOS.MEDIOS[14:15])
sum(poblacionm*residuales1)
cov(poblacionm,residuales1)
sum(I(poblacionm^2)*residuales1)
cov(I(poblacionm^2),residuales1)

#Supuesto 5: Homocedasticidad (Prueba White)
library(stargazer)
regresion_auxiliar<-lm(I(residuales1^2)~I(poblacionm^4)+I(poblacio
nm^3)+I(poblacionm^2)+poblacionm)
resumen<-summary(regresion_auxiliar)
tabla1<-as.data.frame(cbind(residuales1))
n<-nrow(tabla1)
R2<-resumen$r.squared
WHT<-n*R2
gl<-5-1
valorp<-1-pchisq(q=WHT,df=gl)
valorcritico<-qchisq(p=.95,df=gl)
salida.names<-c(WHT, valorcritico,valorp)
names(salida.names)<-c("Estadístico WHT","Valor Crítico", "Valor P")
stargazer(salida.names, title="Resultados de la Prueba de homocedas
ticidad", type="text", digits = 2)

# Usando la libreria "lmtest"
```

```
library(lmtest)
prueba_white<-bptest(regresion_1,~I(poblacionm^4)+I(poblacionm^3)+
I(poblacionm^2)+poblacionm)
print(prueba_white)

#Apoyo gráfico
plot(residuales1,xlim=c(0,length(residuales1)), ylim=c(min(residua
les1)-.01,max(residuales1)+.01), col="blue", pch=19)
abline(h=min(residuales1), lty=2, col="red", lwd=2)
abline(h=max(residuales1), lty=2, col="red", lwd=2)

#Supuesto 6: Evaluacion en la media de x resulta en la media de y.
x_media<-mean(poblacion_media)
y_media<-mean(pendiente)
x2_media<-mean(I(poblacion_media^2))
p<-A*x2_media+B*x_media+D
(p-y_media)
# El resultado debe ser muy cercano a 0 para validar este supuesto.

#Supuesto 7: La covarianza muestral entre los residuales y los valores
ajustados es nula.
valoresajustados1<-regresion_1$fitted.values
sum(residuales1*valoresajustados1)
cov(residuales1,valoresajustados1)

#Supuesto 8: Ausencia de autocorrelación entre los residuales
library(lmtest)
dwtest(regresion_1, alternative="two.sided", iterations=100)

#Supuesto 9: Los residuales siguen una distribución normal
#Gráfico Q-Q
qqnorm(residuales1, pch=19, col="black")
qqline(residuales1,lty=2, col="red")

#Histograma
h1<-hist(residuales1, xlab="RESIDUALES", main="NORMALIDAD DE LOS RESI
DUALES")
x_ajus<-seq(min(residuales1),max(residuales1),length = 200)
y_ajus<-dnorm(x_ajus, mean=mean(residuales1),sd=sd(residuales1))
y_ajus<-y_ajus*diff(h1$mids[1:2])*length(residuales1)
lines(x_ajus,y_ajus, col="blue", lwd=2)

#Prueba Shapiro Test
shapiro.test(residuales1)
```

```
 #Modelo 2: Regresión para la función transformada de la población.

#Supuesto 2: Variabilidad en la variable independiente.
periodo<-c(datos$PERIODO[1:3],datos$PERIODO[5:16])
factor(periodo)

# Supuesto 3
residuales2<- regresion_2$residuals
#La suma de los residuales es 0.
sum(residuales2)
#La media de los residuales es 0.
mean(residuales2)

#Supuesto 4: La covarianza muestral entre los regresores y los resi
duales es 0.
sum(periodo*residuales2)
cov(periodo,residuales2)

#Supuesto 5: Homocedasticidad (Prueba White)
library(stargazer)
regresion_auxiliar2<-lm(I(residuales2^2)~I(periodo^2)+periodo)
resumen2<-summary(regresion_auxiliar2)
tabla2<-as.data.frame(cbind(residuales2))
n2<-nrow(tabla2)
R2.2<-resumen2$r.squared
WHT2<-n2*R2.2
gl2<-3-1
valorp2<-1-pchisq(q=WHT2,df=gl2)
valorcritico2<-qchisq(p=.95,df=gl2)
salida.names2<-c(WHT2, valorcritico2,valorp2)
names(salida.names2)<-c("Estadístico WHT","Valor Crítico", "Valor P")
stargazer(salida.names2, title="Resultados de la Prueba de homocedas
ticidad", type="text", digits = 2)

# Usando la libreria "lmtest"
library(lmtest)
prueba_white2<-bptest(regresion_2,~I(periodo^2)+periodo)
print(prueba_white2)

#Apoyo gráfico
plot(residuales2,xlim=c(0, length(residuales2)), ylim=c(min(residua
les2)-.01,max(residuales2)+.01), col="blue", pch=19)
abline(h=min(residuales2), lty=2, col="red", lwd=2)
abline(h=max(residuales2), lty=2, col="red", lwd=2)
```

```
#Supuesto 6: Evaluacion en la media de x resulta en la media de y.
y_media<-mean(transformada)
x_media<-mean(periodo)
p<-alfa*x_media+beta
(y_media-p) # El resultado debe ser muy cercano a 0 para validar este
supuesto.

#Supuesto 7: La covarianza muestral entre los residuales y los valores
ajustados es nula.
valoresajustados2<-regresion_2$fitted.values
sum(residuales2*valoresajustados2)
cov(residuales2,valoresajustados2)

#Supuesto 8: Ausencia de autocorrelación entre los residuales.
library(lmtest)
dwtest(regresion_2, alternative="two.sided", iterations=100)

#Supuesto 9: Los residuales siguen una distribución normal.
#Gráfico Q-Q
qqnorm(residuales2, pch=19, col="black")
        qqline(residuales2,lty=2, col="red")

#Gráfica distribución normal.
h2<-hist(residuales2,xlab="RESIDUALES", main="NORMALIDAD DE LOS RESI
DUALES")
x_ajus2<-seq(min(residuales2),max(residuales2),length = 300)
y_ajus2<-dnorm(x_ajus2, mean=mean(residuales2),sd=sd(residuales2))
y_ajus2<-y_ajus2*diff(h2$mids[1:2])*length(residuales2)
lines(x_ajus2,y_ajus2, col="red", lwd=2)

#Prueba Shapiro- Wilk
shapiro.test(residuales2)
```

Bibliografía

Almeka, L., Gerland, P., Raftery, A. & Wilmoth, J. (2015). *The United Nations Probabilistic Projections: An introduction to demographic forecasting with uncertainly. Recuperado el 29 de marzo de 2020.* https://www.ncbi.nlm.nih.gov/pmc/articles/PM46%2062414/

Ambientum. (2018). *Teoría malthusiana: crecimiento mundial de la población.* https://www.ambientum.com/ambientum/agricultura/teoria-malthusiana-crecimiento-mundial-de-la-poblacion.asp

Barrios, E., Garcia-J. & Matuk-J. (2016). *Tablas de probabilidad.* México, Instituto Tecnológico Autónomo de México.

Beaumont, A. & Pierce. (1963). *The Algebraic Foundations of Mathematics.* United States of America, Addison Wesley Publishing Company Inc.

Beers, B. (2020). *P-Value.Recuperado el 12 de abril de 2020.* https://www.investop%20edia.com/terms/p/p-value.asp

Bravo, A. (2017). *Ecuaciones Diferenciales: un enfoque moderno.* Perú, Universidad Alas Peruanas.

Canavos, G. (1988). *Probalidad y Estadística. Aplicaciones y métodos.* México, Mc Graw Hill.

Carter, R., Griffiths-W. & Lim-G. (2011). *Principles of econometrics.* USA, WILEY.

CEDALE. (2019). *Tendencias recientes de la población de América Latina y el Caribe.Recuperado el 22 de abril de 2020.* https://www.cepal.org/sites/default/files/static/files/dia_mundial_de_la_poblacion_2019.pdf

Colegio-de-México. (2012). *Cálculo Integral.Recuperado el 10 de enero de 2020.* https://www.studocu.com/es-mx/document/el-colegio-de-mexico/%20matematicas-i/apuntes/calculo-integral/3755503/view

CONAPO. (2016). *Proyecciones de la población de México y de las entidades federativas 2016-2050.Recuperado el 27 de abril de 2020.* https://www.gob.mx/cms/uploads/attachment/file/487382/21_PUE.pdf

CONAPO. (2019). *Proyecciones de la población de los municipios de México, 2015-2030.Recuperado el 13 de marzo de 2020.* https://www.gob.mx/conapo/documentos/proyecciones-de-la-poblacion-de-los-municipios-de-mexico-2015-2030?idiom=es

Del-Pino, C. (2019). *Métodos de integración por fracciones parciales.Recuperado el 19 de abril de 2020.* http://matesup.cl/calc2/unidad1/AP

Dietrichson, A. (2019). *Métodos Cuantitativos.Recuperado el 5 de marzo de 2020.* https://bookdown.org/dietrichson/metodos-cuantitativos/test-de-normalidad.html

Enriquez, J. (2019). *11 de julio - Día Mundial de la Población.Recuperado el 10 de octubre de 2020.* http://www.udg.mx/es/efemerides/11-de-julio-dia-mundial-de-la-poblacion

FAO. (2006). *La ganadería amenaza el medio ambiente.* http://www.fao.org/newsroom/es/news/2006/1000448/index.html

Georgia-Institute-of-Technology. (2014). *The Logistic Equation. Recuperado el 19 de febrero de 2020.* https://www.coursehero.com/file/2085%201271/logisticgrowth

González-Rosas, J. & Zárate-Gutiérrez, I. (2018). *The Stable Bounded Theory an Alternative to Projecting Populations. The Case of Mexico. Recuperado el 10 de febrero de 2020.* https://journalofbusiness.org/index.php/GJMBR/article/view/2570

Gujarati, D. & Porter-D. (2010). *Econometría.* México, McGraw-Hill.

Hernández, M., López-R & Velarde-S. (2013). *La situación demográfica en México. Panorama desde las proyecciones de población.* México, CONAPO.

INEGI. (2009). *Mujeres y hombres de Puebla.Recuperado el 24 de abril de 2020.* http://cedoc.inmujeres.gob.mx/ftpg/Puebla/Muj_Puebla.pdf

INEGI. (2015). *Censos y conteos de población y viviendaRecuperado el 4 de febrero de 2020.* web:https://www.inegi.org.mx/programas/ccpv/1895/default.html#Tabulados

INEGI. (2019). *Estadísticas a propósito del día mundial de la población.Recuperado el 2 de mayo de 2020.* https://www.inegi.org.mx/contenidos/salade%20prensa/aproposito/2019/Poblacion2019_Nal.pdf

Infante, S. & Zarate-G. (1990). *Métodos Estadísticos. Un enfoque interdisciplinario.* México, Editorial Trillas.

Kenton, W. (2019). *Durbin-Watson statistic.Recuperado el 8 de abril de 2020.* https://www.investopedia.com/terms/d/durbin-watson-statistic.asp

Kyun, T. (2015). *T test as a parametric statistic.* Busan,Korea, Korean Journal of Anesthesiology.

Lomborg, B. (2014). *El problema demográfico.* https://www.nacion.com/opinion/foros/el-problema-demografico/W7LMNYLH5ZCARF6BOFVX423UBI/story/

López, V. (2019). *Puebla, en proceso de colapso urbano por crecimiento poblacional.Recuperado el 22 de abril de 2020.* https://www.milenio.com/politica/comunidad/puebla-en-proceso-de-colapsourbano-por-creci%20miento-poblacional

Malthus, T. (1826). *An essay on the principle of population.* London, Albermale Street.

Medhi, J. (1981). *Stochastic Processes.* New York, WILEY.

Mendoza, A., Reyes-H & González-J. (2020). *Esperanza de vida en el estado de Puebla: Pronósticos bajo la teoría de estabilidad acotada. Recuperado el 06 de marzo de 2020.* http://bit.ly/2vYYGXM

Montgomery, D., Peck-E. & Vining-G. (2012). *Introduction to linear regression.* USA, WILEY.

Novales, A. (2010). *Análisis de Regresión.* Madrid, Universidad Complutense.

ONU. (2017). *The Cohort Component Method for Making Population Projections.Recuperado el 29 de marzo de 2020.* http://www.un.org/esa/population/techcoop/PopProj/module1%20/chapter2.pdf

OpenStax. (2018). *Population Growth and Regulation.* https://cnx.org/contents/X6dCGi4e@12/Crecimiento-y-regulaci%C3%B3n-de-la-poblaci%C3%B3n

Parsons, J. (1991). *Population Control and Politics.* United States of America, Springer.

Perez-J. (2008). *Integral de Riemann.* España, Universidad de Granada.

Population-Reference-Bureau. (2017). *Understanding and using Population projections.Recuperado el 27 de abril de 2020.* http://www.prb.org/Publications/Reports/%202021/UnderstandingandUsingPopulation.aspx

R-STUDIO. (2019). *Programa R.Recuperado el 30 de abril de 2020.* rstudio.com

Ruiz, J. (2011). La transición demográfica y el envejecimiento poblacional: Futuros retos para la política de salud en México. *Encrucijada*, (8), 1-16.

Tuirán, R. (1988). *La situación demográfica en México.* https://www.redalyc.org/articulo.oa?id=11201603

University-of-Toronto. (2013). *Linear Regression.Recuperado el 24 de febrero de 2020.* https://www.coursehero.com/file/25907746/lecture3pdf/%20UNTE-31-38.pdf

Villa-D. (2019). *Estimación de los recursos monetarios del ISSTE para la salud de sus derechohabientes de la tercera edad.* Puebla, BUAP.

Weinstein, J. & Pillai-V. (2016). *Demography: The Science of Population.* USA, Rowman Littlefield.

Wooldridge, J. (2009). *Introducción a la econometría. Un enfoque moderno.* USA, South-Western Cengage Learning.

Zill, G. (2011). *A First Course In Differential Equations with Modeling Applications.* Boston, Brooks/Cole.