

**UNIVERSIDAD DE EL SALVADOR
FACULTAD DE CIENCIAS Y HUMANIDADES
ESCUELA DE POSGRADOS**



**TITULO:
“VIOLENCIA HOMICIDA EN EL SALVADOR”
UNA MIRADA DESDE LA GOBERNANZA**

PRESENTADO POR:

LICENCIADA ANA BEATRIZ VALDEZ ÁLVAREZ (VA 98029)
LICENCIADO WALTER ALEXANDER PÉREZ CRUZ (PC 98005)

**TRABAJO FINAL PARA OPTAR AL TÍTULO DE:
MAESTRA/O EN MÉTODOS Y TÉCNICAS DE INVESTIGACIÓN SOCIAL**

ASESOR:

DOCTOR RODRIGO CRUZ ORELLANA LEÓN

**CIUDAD UNIVERSITARIA, DR. FABIO CASTILLO FIGUEROA, SAN
SALVADOR CENTRO, EL SALVADOR, CENTROAMÉRICA, JULIO DEL 2024**

AUTORIDADES DE LA UNIVERSIDAD DE EL SALVADOR

INGENIERO JUAN ROSA QUINTANILLA QUINTANILLA

RECTOR

DOCTORA EVELYN BEATRIZ FARFÁN

VICERRECTORA ACADÉMICO

MAESTRO RÓGER ARMANDO ARIAS ALVARADO

VICERRECTOR ADMINISTRATIVO

LICENCIADO PEDRO ROSALÍO ESCOBAR CASTANEDA

SECRETARIO GENERAL

LICENCIADA ANA RUTH AVELAR

DEFENSORA DE LOS DERECHOS UNIVERSITARIOS

LICENCIADO CARLOS AMÍLCAR SERRANO RIVERA

FISCAL GENERAL

AUTORIDADES DE LA FACULTAD DE CIENCIAS Y HUMANIDADES

MAESTRO JULIO CÉSAR GRANDE RIVERA
DECANO

MAESTRA MARÍA BLAS CRUZ JURADO
VICEDECANA

MAESTRA NATIVIDAD TESHÉ PADILLA
SECRETARIA

MAESTRA SANDRA LORENA BENAVIDES DE SERRANO
DIRECTORA ESCUELA DE POSGRADO

DOCTORA XIOMARA DEL CARMEN AVENDAÑO
COORDINADORA DE PROCESOS DE GRADO

DOCTORA ANGELA JEANNETTE AURORA ANAYA
COORDINADORA DEL PROGRAMA

Agradecimientos

Primeramente, dar las gracias a Dios por permitirme culminar esta investigación, sin él no hubiera sido posible este logro.

A mis padres Miguel Valdez y Francisca Álvarez de Valdez; a mis hijos, Ester Beatriz Lemus Valdez, Gerber Rolando Lemus Valdez y su padre, Rolando Lemus Gómez† por su comprensión durante este recorrido y su apoyo que ha sido una fortaleza para seguir adelante.

A mi asesor de tesis, el Dr. Cruz Rodrigo Orellana León, por adquirir el compromiso de orientar este tema de investigación que con un buen criterio y trabajo en equipo se ha logrado concluir.

A mi compañero de tesis Walter Alexander Pérez Cruz, por su confianza, dedicación perseverancia para llegar al final de este proceso.

Y a todas aquellas personas que de una u otra forma colaboraron con opiniones y observaciones necesarias para poder dar por finalizado satisfactoriamente este largo, pero fructífero camino en la realización de este tema investigación.

Ana Beatriz Valdez Álvarez

Agradecimientos

Quiero dedicar este logro a aquellos seres queridos que ya no están en el plano terrenal y a dos de ellos en particular, que hace nueve años me guardan y me siguen inspirando desde el cielo, para ellos mis agradecimientos eternos.

Dios ha sido un pilar fundamental en todo este tiempo, por lo cual estoy muy agradecido por sus bondades y misericordias.

A mi querida hija Kelly Raquel y a mi pequeña bebé Ana Victoria, quienes son dos luceros que me impulsan día a día, dedico también este logro.

Agradezco enormemente a mi compañera de tesis Ana Beatriz Valdez Álvarez, quien con su capacidad e intelecto ha sido fundamental para la realización de esta investigación.

También quiero expresar un enorme agradecimiento al Dr. Cruz Rodrigo Orellana León, quien asumió el reto de acompañarnos y asesorarnos a lo largo de este trabajo, y que oportunamente nos orientó con su conocimiento.

A todas aquellas personas que de una u otra forma contribuyeron con sus apoyos, consejos, orientaciones y ánimos para culminar esta maestría, quiero expresarles mis más sinceros agradecimientos.

Walter Alexander Pérez Cruz

Índice General

Contenido	Página
Índice General	vi
Índice de figuras	ix
Índice de tablas	x
Resumen	xiii
Introducción	xiv
Capítulo 1. Planteamiento del problema	13
1.1 Delimitación del problema	13
1.2 Objetivos	15
1.2.1 Objetivo general	15
1.2.2 Objetivos específicos	15
1.3 Justificación	16
Capítulo 2. Fundamentos teóricos de la investigación: sociales y estadísticos	18
2.1 Fundamentos teóricos sociales	18
2.1.1 Definición de violencia	22
2.1.2 Tipos de violencia	24
2.1.3 Variables que explican la violencia	27
2.1.3.1 Pobreza	27
2.1.3.2 Desigualdad	27
2.1.3.3 Desempleo	28
2.1.3.4 Producto Interno Bruto per cápita (PIB)	28
2.1.3.5 Índice de Desarrollo Humano (IDH)	29
2.1.3.6 Institucionalidad o Gobierno de La Ley (Rule of Law)	29
2.1.3.7 Violencia expresada en homicidios	30
2.2 Fundamentos teóricos estadísticos	30
2.2.1 Regresión Lineal Simple	30
2.2.1.1 Análisis univariado	30
2.2.1.2 Análisis bivariado	30
2.2.1.3 Regresión lineal simple	31
2.2.1.4 Diagrama de dispersión	31

2.2.1.5 Coeficiente de correlación	33
2.2.1.6 Coeficiente de determinación	33
2.2.1.7 Análisis de la varianza para regresión lineal	34
2.2.1.8 Distribución F	36
2.2.2 Regresión Lineal Múltiple	38
2.2.2.1 Normalidad	38
2.2.2.2 Linealidad	39
2.2.2.3 Independencia	41
2.2.2.4 Homocedasticidad	41
2.2.2.5 Colinealidad	41
2.2.2.6 Construcción de los modelos	43
2.2.3 Métodos de selección de modelos	44
2.2.3.1 Criterio de Información de Akaike	44
2.2.3.2 Criterio de Información Bayesiana (BIC)	45
2.2.3.3 Estadístico Predictivo Conceptual (Cp) de Mallows	45
2.2.4 Modelos lineales generalizados: distribución de Poisson	46
2.2.4.1 Función logaritmo	46
Capítulo 3. Fundamentación Metodológica	51
3.1 Enfoque y método de investigación	51
3.1.1 Enfoque	51
3.1.2 Método	52
3.2 Tipo de Investigación	53
3.2.1 Alcance de la investigación	53
3.2.2 Diseño de la investigación	53
3.3 Población y Muestra	54
3.4 Técnica e instrumento de recolección de datos	55
3.5 Procesamiento y análisis de los datos	55
3.6 Operacionalización de las variables	58
Capítulo 4. Resultados	59
4.1 Análisis exploratorio de los datos	59
4.1.1 Análisis univariado	59

4.1.2 Análisis bivariado	71
4.1.2.1 Coeficiente de correlación.	71
4.1.2.2 Coeficiente de determinación.	73
4.2 Regresión Lineal Múltiple	74
4.2.1 Normalidad de las series	75
4.2.2 Linealidad	76
4.2.3 Independencia	77
4.2.4 Homocedasticidad	78
4.2.5 Multicolinealidad	79
4.3. Modelo Lineal Múltiple	81
4.3.1 Selección de los modelos	81
4.3.2 Selección de los modelos bajo los criterios Cp, AIC y BIC	85
4.3.3 Propuesta de los Modelos Lineales	87
4.4 Modelo lineal generalizado	89
4.4.1 Modelo de Regresión de Poisson	89
4.4.2 Interpretación de los resultados	91
Capítulo 5. Discusión	111
Conclusiones	121
Recomendaciones	123
Referencias	125
Anexo 1. Base de datos	131
Anexo 2. Prueba de los 10 modelos	132
Anexo 3. Cálculo de Error Cuadrático Medio	135

Índice de figuras

Figura 1 Diagrama de Dispersión Lineal	32
Figura 2 Distribución F y la Región de Rechazo de la Hipótesis	37
Figura 3 Representación Gráfica del Comportamiento de las Variables en Estudio en el Periodo 1990 a 2019	68
Figura 4 Gráfica del Comportamiento de la Variable Homicidios en el Periodo 1990 a 2019	70
Figura 5 Comportamiento de Valores Predichos y Observados para el Modelo 1	92
Figura 6 Comportamiento de Valores Predichos y Observados para el Modelo 2	94
Figura 7 Comportamiento de Valores Predichos y Observados para el Modelo 3	96
Figura 8 Comportamiento de Valores Predichos y Observados para el Modelo 4	98
Figura 9 Comportamiento de Valores Predichos y Observados para el Modelo 5	100
Figura 10 Comportamiento de Valores Predichos y Observados para el Modelo 6	101
Figura 11 Comportamiento de Valores Predichos y Observados para el Modelo 7	103
Figura 12 Comportamiento de Valores Predichos y Observados para el Modelo 8	104
Figura 13 Comportamiento de Valores Predichos y Observados para el Modelo 9	106
Figura 14 Comportamiento de Valores Predichos y Observados para el Modelo 10	108

Índice de tablas

Tabla 1 Fórmulas para el análisis de varianza (ANOVA)	36
Tabla 2 Rangos en los que puede variar el coeficiente de correlación de Spearman	40
Tabla 3 Operacionalización de las variables	58
Tabla 4 Estadísticos descriptivos de las variables en estudio en el periodo 1990 a 2019	67
Tabla 5 Estadísticos descriptivos de la variable homicidios en el periodo 1990 a 2019	69
Tabla 6 Correlaciones de Pearson entre las variables explicativas y los homicidios	72
Tabla 7 Coeficiente de determinación entre las variables en estudio	73
Tabla 8 Resultados del test de Shapiro -Wilk para analizar normalidad en la serie de datos	75
Tabla 9 Resultado del test de Spearman para analizar correlación en las series de datos	76
Tabla 10 Estadística de contraste del test de Spearman para analizar correlaciones lineales	77
Tabla 11 Estadísticas del test de Spearman para analizar independencia	78
Tabla 12 Resultado del análisis de la varianza de cada serie de datos	79
Tabla 13 Resultado del análisis de correlación de Rho de Spearman	80
Tabla 14 Construcción de las 63 combinaciones posibles entre las variables en estudio	82
Tabla 15 Selección de modelos de acuerdo a los criterios CP, AIC Y BIC	86
Tabla 16 Posibles modelos seleccionados mediante el criterio CP, AIC Y BIC	87
Tabla 17 Modelo 1. Parámetros del modelo compuesto por la variable homicidios (h) y la variable pobreza (pobext)	91
Tabla 18 Modelo 2. Parámetros del modelo compuesto por la variable homicidios (h) y las variables GINI Y PIB	93
Tabla 19 Modelo 3. Parámetros del modelo compuesto por la variable homicidios (h) y las variables GINI y pobreza	95
Tabla 20 Modelo 4. Parámetros del modelo compuesto por la variable homicidios (h) y las variables GINI, desem (desempleo) y pobext (pobreza)	97
Tabla 21 Modelo 5. Parámetros del modelo compuesto por la variable homicidios (h) y las variables GINI, IDH y pobext (pobreza)	99
Tabla 22 Modelo 6. Parámetros del modelo compuesto por la variable homicidios (h) y las variables GINI, desem (desempleo), IDH y pobext (pobreza)	101
Tabla 23 Modelo 7. Parámetros del modelo compuesto por la variable homicidios (h) y las variables GINI, PIB, pobext (pobreza) y G&L (gobierno de la ley)	102

Tabla 24 Modelo 8. Parámetros del modelo compuesto por la variable homicidios (h) y las variables GINI, desem (desempleo), PIB, IDH y G&L (gobierno de la ley)	104
Tabla 25 Modelo 9. Parámetros del modelo compuesto por la variable homicidios (h) y las variables GINI, desem (desempleo), IDH, pobext (pobreza) y G&L (gobierno de la ley)	105
Tabla 26 Modelo 10. Parámetros del modelo compuesto por la variable homicidios (h) y las variables GINI, desem (desempleo), PIB, IDH, pobext (pobreza) y G&L (gobierno de la ley)	107

Siglas y Acrónimos

AIC: Criterio de Información de Akaike.

ANOVA: Análisis de Varianza.

BCR: Banco Central de Reserva.

BIC: Criterio de Información Bayesiano.

BM: Banco Mundial.

CBA: Canasta Básica Alimentaria.

CEPAL: Comisión Económica para América Latina y El Caribe.

DESEM: Desempleo.

DIGESTYC: Dirección General de Estadísticas y Censos.

EHPM: Encuesta de Hogares de Propósitos Múltiples.

FESPAD: Fundación de Estudios para la Aplicación del Derecho.

FLACSO: Facultad Latinoamericana de Ciencias Sociales.

G&L: Gobierno de la Ley.

GINI: Índice de GINI.

H: Homicidios.

IDH: Índice de Desarrollo Humano.

IDM: Indicadores del Desarrollo Mundial.

INBI: Indicador de Necesidades Básicas Insatisfechas.

MLG: Modelos Lineales Generalizados.

MRP: Modelo de Regresión de Poisson.

OMS: Organización Mundial de la Salud.

OIT: Organización Internacional del Trabajo.

OPS: Organización Panamericana de la Salud.

PEA: Población Económica Activa.

PIB: Producto Interno Bruto.

PNUD: Programa de las Naciones Unidas para el Desarrollo.

POBEXT: Pobreza extrema.

UNODC: Oficina de las Naciones Unidas contra la Droga y el Delito.

USAID: Agencia de los Estados Unidos para el Desarrollo Internacional.

WGI: Worldwide Governance Indicators.

Resumen

La investigación aborda la problemática de la violencia, específicamente la violencia homicida desde una perspectiva cuantitativa. La propuesta de esta investigación es la construcción de un modelo matemático que contribuya a la explicación de los homicidios en El Salvador, a partir del análisis de un conjunto de seis variables y observar el comportamiento estadístico de estas en relación con la tasa de homicidios a lo largo de 30 años, considerando el periodo de 1990 a 2019. Las variables que se han considerado para el análisis del comportamiento de los Homicidios en El Salvador son: Pobreza extrema, Desigualdad (GINI), Desempleo, Producto Interno Bruto per cápita, Índice de Desarrollo Humano e Institucionalidad. Esta última variable, se incorpora como un elemento nuevo al análisis estadístico de los Homicidios en El Salvador, y tiene que ver con la capacidad que tienen las instituciones que ejercen justicia en el país (Policía Nacional Civil, Fiscalía y los Juzgados). Esta variable ha sido tomada de los indicadores de gobernanza del Banco Mundial (Worldwide Governance Indicators). En esta investigación se realiza una exploración estadística de estas variables y sus correlaciones con la variable Homicidio, dando como resultado la propuesta de un modelo estadístico, permitiendo analizar el comportamiento de los Homicidios a través de las variables que más se ajustan a las pruebas estadísticas del modelo propuesto. Sin embargo, determinadas variables presentan características estadísticas que no cumplen con los supuestos de normalidad. Por lo que fue necesario la aplicación de Modelos Lineales Generalizados (MLG), permitiendo que el modelo final sea un modelo lineal que se encuentre relacionado con la variable dependiente a través de una función de enlace, y la varianza de cada medición sea una función de su valor predicho. Los MLG son aplicados cuando no se cumple con el supuesto de normalidad entre los datos y entre los MLG disponibles, el más apropiado para esta investigación fue el Modelo de Poisson, dado que la variable dependiente es discreta. Finalmente, el modelo obtenido está compuesto por 2 variables: Desigualdad medida a través del índice de GINI y el Producto Interno Bruto (PIB) per cápita.

Palabras Claves: El Salvador, homicidios, gobierno de la ley, modelos lineales generalizados, Poisson.

Introducción

El tema central de esta investigación es la violencia homicida en El Salvador vista desde la Gobernanza, en la cual se plantea un análisis estadístico de seis variables a lo largo del periodo que comprende de 1990 a 2019 que oriente la construcción de un modelo matemático que contribuya a la explicación de los homicidios en El Salvador. Las variables en estudio son: Pobreza extrema, Desigualdad, Desempleo, Producto Interno Bruto, Índice de Desarrollo Humano e Institucionalidad (Gobierno de la Ley), esta última se agrega como un elemento nuevo a la explicación estadística de la violencia homicida.

Este trabajo se encuentra estructurado por cinco capítulos. En el primer capítulo se hace un abordaje de la problemática en la que se deja claro que la violencia, enfocada en los homicidios, en el caso de El Salvador, requiere de la atención científica para su abordaje, debido a las altas tasas que ha presentado por años. También se explica la necesidad de incorporar nuevas variables al análisis de la variación en la tasa de homicidios, como es la institucionalidad, en vista de que las variables que tradicionalmente se han utilizado para explicar la violencia, a lo largo del periodo en estudio, han tenido mejoras en sus indicadores.

En el segundo capítulo se realiza una explicación teórica de la violencia y sus tipos, poniendo énfasis en el sistema normativo que regula el comportamiento de los seres humanos. De igual forma se explican teóricamente las variables que han sido utilizadas para explicar el fenómeno de la violencia en El Salvador. Dentro de este mismo capítulo, se realiza la fundamentación teórica de los pasos y procedimientos estadísticos que se han aplicado en la investigación.

El tercer capítulo, se refiere al proceso metodológico llevado a cabo con el fin de explicar paso a paso los procedimientos que se siguieron para la construcción del modelo estadístico. El estudio es eminentemente cuantitativo, con un enfoque hipotético de carácter estadístico. El nivel de estudio es correlacional, pero pasa por el descriptivo, hasta un nivel predictivo que concluye con la elección de un modelo estadístico, que determinó, del conjunto de las seis variables, cuáles de estas tienen mayor significancia, en la incidencia de los homicidios como variable dependiente.

En el cuarto capítulo, se presentan los resultados, comenzando con un análisis exploratorio y descriptivo, seguido de un análisis bivariado para identificar relaciones de

dependencia entre variables. Este proceso culmina en la construcción de un modelo de regresión múltiple que explica el comportamiento de la violencia homicida en El Salvador. Además, se discute la aplicación de Modelos Lineales Generalizados (MLG), seleccionando el modelo más adecuado para esta investigación. El Modelo de Poisson fue elegido debido a que la variable dependiente, homicidios, es discreta, lo que implica que su medición se realiza en datos enteros.

El último capítulo, aborda la discusión de los resultados, destacando que el conjunto de variables seleccionadas para esta investigación (variables independientes) presentan algún nivel de correlación con la variable homicidios. Dentro de estas la que mayor correlación presenta es la Pobreza extrema. Se concluye con la explicación del modelo compuesto por 2 variables: Desigualdad (GINI) y Producto Interno Bruto (PIB) per cápita, como el que mejor explica la variable homicidio.

Capítulo 1. Planteamiento del problema

Abordar la problemática de la violencia expresada a través de los homicidios, es esencial para comprender y abordar los desafíos sociales y de seguridad que enfrentan las comunidades en todo el mundo. Este fenómeno, que afecta a individuos, familias y comunidades enteras, tiene raíces profundas y complejas que requieren un análisis multidisciplinario para su comprensión y prevención efectivas.

En el caso de El Salvador resalta de manera significativa esta forma de violencia, evidenciando la necesidad imperante de llevar a cabo análisis exhaustivos y profundos sobre este fenómeno. La situación en El Salvador ha sido objeto de atención a nivel nacional e internacional debido a los altos índices de violencia, especialmente los relacionados con homicidios.

La ciencia estadística y la ciencia social desempeñan roles cruciales en la comprensión de la violencia homicida y en la identificación de factores asociados que pueden contribuir a su ocurrencia. Desde una perspectiva estadística, se emplean técnicas y herramientas para recopilar, analizar e interpretar datos relacionados con los homicidios.

Por otro lado, la ciencia social aporta un marco teórico y conceptual para comprender la naturaleza y dinámica de la violencia homicida. Se investigan cómo la desigualdad socioeconómica, la falta de oportunidades educativas y laborales, la exposición a la violencia en el hogar o la comunidad, la presencia de armas de fuego y la eficacia de las políticas de seguridad pública, entre otros explican la violencia en las sociedades.

La colaboración entre la ciencia estadística y la ciencia social permite una comprensión más completa y holística de la violencia homicida. Los análisis estadísticos pueden identificar patrones y tendencias en los datos, mientras que los enfoques teóricos y metodológicos de la ciencia social proporcionan un contexto significativo para interpretar estos hallazgos.

1.1 Delimitación del problema

La violencia en todas sus manifestaciones ha sido una triste realidad que ha permeado de manera generalizada en América Latina. Sin embargo, es importante destacar que esta problemática no se presenta con la misma intensidad en todos los países de la región, ya que su prevalencia está influenciada por una compleja interacción de factores sociales, económicos, políticos y culturales.

Entre las diversas formas de violencia, resulta especialmente alarmante el incremento de la violencia homicida en países como México, Colombia, Brasil, Venezuela y El Salvador, donde se ha observado un notable aumento en las tasas de homicidios en las últimas décadas (Briceño-León, 2012).

El Salvador, se ha posicionado entre los países de América Latina con uno de los índices más altos de muertes debido a actos violentos. Las estadísticas muestran tasas récord de homicidios que alcanzan aproximadamente los 70 por cada 100,000 habitantes, una cifra alarmante que refleja la grave situación de violencia que enfrenta la población salvadoreña. Dentro de este contexto, resulta especialmente preocupante el perfil de las víctimas, siendo los hombres de entre 18 y 30 años el grupo demográfico más afectado por estos crímenes (PNUD, 2013).

Esta realidad no solo representa una pérdida irreparable de vidas humanas, sino que también genera un profundo impacto en la sociedad salvadoreña, afectando la estabilidad y el desarrollo del país.

Para algunos autores, como González-Pérez et al. (2012) el fenómeno de la violencia se ha convertido en un “problema de salud pública”, puesto que las tasas de homicidios son consideradas epidémicas, por lo que el fenómeno se ha convertido en objeto de análisis por parte de las autoridades de salud pública. Por su parte, Pierre Salama (2012) habla de la violencia como violencia extrema cuando esta llega a los 100 homicidios por cada 100 mil habitantes.

Una explicación inmediata es la vinculación con la pobreza, se dice que la violencia se usa de manera funcional, es decir, el individuo que tiene la falta de un bien usa la violencia para obtener su meta que es el bien deseado. La violencia en nuestras realidades latinoamericanas lleva, en ocasiones, a emitir juicios de coexistencia; según Buvinic et al. (1999) no hay demostraciones empíricas que realmente satisfagan en torno a la generación de una base para la existencia de la otra, es decir que no puede afirmarse que únicamente la pobreza genera violencia.

Otra explicación ha sido la desigualdad, cabe resaltar que América Latina es considerada como la región más desigual del mundo (Brender, 2012; Buvinic et al., 2005; Rodríguez, 2004). La pobreza que se genera a raíz de la desigualdad en los ingresos, en el acceso a oportunidades y recursos, presenta mayores razones explicativas al surgimiento de la violencia (Fajnzlber, Lederman y Loayza, 2002). Es esta desigualdad la que favorece también la permanencia de la

violencia en distintos escenarios sociales aunado a que, según el Banco Interamericano de Desarrollo (1998) es la desigualdad el primer factor estructural de mayor riesgo para la violencia.

Sin embargo, el comportamiento que ha descrito en los últimos años el crecimiento de la tasa de homicidios, frente a indicadores que muestran una alentadora reducción de la pobreza y de la desigualdad (Véase los indicadores de la Encuesta de Hogares de Propósitos Múltiples (EHPM) de la Dirección General de Estadísticas y Censos, DIGESTYC), así como el mejoramiento y aumento de indicadores de desarrollo humano (Véase los indicadores del Programa de las Naciones Unidas para el Desarrollo, PNUD), podría indicar que la violencia no se debe únicamente a la desigualdad o la pobreza, como se ha explicado tradicionalmente.

En este sentido, se busca incorporar al análisis y discusión que se ha realizado de la violencia homicida en El Salvador, el indicador del Gobierno de la Ley, el cual hace referencia a la institucionalidad que se “relaciona con el desempeño de las autoridades judiciales y penales, con la legitimidad de las normas y leyes, las condiciones políticas y el pacto social” (Briceño-León et al., 2012, p. 38). Este indicador puede encontrarse en Worldwide Governance Indicators del Banco Mundial.

1.2 Objetivos

1.2.1 Objetivo general

Construir una función matemática que exprese lo mejor posible el tipo de relación entre la variación de la tasa de homicidios y la Pobreza extrema, Desigualdad, Desempleo, Producto Interno Bruto, Índice de Desarrollo Humano e Institucionalidad en El Salvador durante el periodo 1990 – 2019.

1.2.2 Objetivos específicos

- Describir las variables Pobreza extrema, Desigualdad, Desempleo, Producto Interno Bruto, Índice de Desarrollo Humano e Institucionalidad que inciden en la tasa de homicidios.
- Explorar los datos de las variables Pobreza extrema, Desigualdad, Desempleo, Producto Interno Bruto, Índice de Desarrollo Humano e Institucionalidad que inciden en la tasa de homicidios.

- Relacionar los homicidios con la pobreza extrema, la desigualdad, el desempleo, el producto interno bruto, el índice de desarrollo humano y la institucionalidad.
- Comprobar mediante pruebas estadísticas las variables que explican las variaciones en la tasa de homicidios, entre Pobreza extrema, Desigualdad, Desempleo, Producto Interno Bruto, Índice de Desarrollo Humano e Institucionalidad.
- Determinar el tipo de modelo matemático que mejor se ajuste para explicar la variación de los homicidios a partir de las variables seleccionadas.

1.3 Justificación

El Salvador, está posicionado como uno de los países de América Latina con mayor índice de muertes por actos violentos, con tasas récord de homicidios que, antes de marzo¹ de 2012, rondaban los 70 homicidios por cada 100,000 habitantes (PNUD, 2013).

Este fenómeno de violencia de El Salvador atrajo la atención de investigadores quienes han buscado explicaciones desde diferentes áreas del conocimiento: económicas, sociales, culturales y políticas. En buena parte estas explicaciones que han caracterizado el estudio de la violencia homicida en El Salvador han estado asociadas principalmente a la pobreza y la desigualdad social (Carcach et al., 2011; Valiente, 2019), estas explicaciones parten de las marcadas desigualdades sociales y económicas que han estado presentes a lo largo de la historia de El Salvador. Sin embargo, no se pueden considerar como las únicas explicaciones, puesto que en los últimos años las tasas de pobreza han mostrado mejorías, al igual que la desigualdad y el desarrollo humano, esto lleva a considerar la búsqueda de otras alternativas teóricas y metodológicas que permitan explicar el comportamiento y aumento de la tasa de homicidios en el país.

De ahí que esta investigación surja por la necesidad de buscar nuevas explicaciones al aumento de los homicidios en El Salvador en un contexto en el que los indicadores de pobreza y desigualdad han mostrado mejorías para la población.

Así, se propone la incorporación de un nuevo elemento a la interpretación de la violencia homicida en El Salvador, a partir de la evidencia que se encuentre en las pruebas, análisis e

¹ En marzo de 2012 se dio inicio lo que se denominó tregua entre pandillas, apoyada por el Gobierno de El Salvador, la cual duraría hasta enero de 2015. Esta tregua redujo significativamente el número de homicidios en el país.

interpretación estadística, tanto de las variables utilizadas tradicionalmente para explicar la violencia, como la incorporación del indicador: estado de derecho o gobierno de la ley (the rule and law).

El concepto de estado de derecho se refiere a la idea de un sistema jurídico en el cual todas las personas, instituciones y entidades, públicas y privadas, incluido el propio Estado, están sometidas al imperio de la ley. Esto implica que las leyes son claras, públicas, aplicadas de manera justa y eficiente, y que existen mecanismos para garantizar su cumplimiento y para proteger los derechos de los ciudadanos.

En el contexto de El Salvador, el estado de derecho ha sido objeto de debate y preocupación debido a la percepción de que las instituciones encargadas de garantizarlo, como el sistema judicial y las fuerzas de seguridad, no siempre funcionan de manera eficiente y justa. La corrupción, la impunidad y la falta de recursos son algunos de los desafíos que enfrenta el país, entre otros.

En este sentido, se recopilaron y analizaron datos sobre homicidios en El Salvador, así como algunas variables tradicionales utilizadas para explicar la violencia homicida, como la pobreza, el desempleo, la desigualdad, el producto interno bruto y desarrollo humano. Además, se incorpora el indicador de estado de derecho, utilizando datos de organismos internacionales que evalúan la eficiencia de la institucionalidad en el país.

Se espera que los resultados de esta investigación contribuyan a una mejor comprensión de las causas y dinámicas de la violencia homicida en El Salvador, incorporando el estado de derecho como una variable relevante en el análisis. La inclusión del indicador de estado de derecho puede permitir identificar posibles relaciones entre la eficiencia de la institucionalidad y la incidencia de la violencia homicida, como ya se hecho para otros países con índices de violencia como los de El Salvador.

Capítulo 2. Fundamentos teóricos de la investigación: sociales y estadísticos

Para profundizar en el fenómeno de la violencia es importante establecer algunos elementos teóricos previos, tal como lo plantea Sauto (2005) al decir que la teoría es acumulativa:

Que siempre comienza por el conocimiento previo. Más aún, los resultados de variados enfoques nutren y son útiles a todas las disciplinas. El conocimiento científico siempre es discutible y provisorio, por lo cual requiere para su crítica que se hagan explícitos las teorías y los métodos utilizados. (p. 24)

Esta perspectiva subraya la importancia de establecer una base teórica sólida que oriente el análisis y la interpretación del fenómeno en estudio. En el caso específico de la violencia homicida, esta aproximación teórica adquiere una relevancia particular, ya que se trata de un fenómeno complejo que se ha definido desde varias disciplinas del saber.

De igual forma establecer la base teórica estadística para la construcción en última instancia de un modelo estadístico, que permita una aproximación a la comprensión de un fenómeno complejo como lo es la violencia homicida y sus determinantes.

Para abordar el análisis de una problemática tan compleja como lo es la violencia, resulta fundamental comenzar con algunas definiciones claves que permitan comprender sus diversas manifestaciones.

2.1 Fundamentos teóricos sociales

El abordaje teórico social para la comprensión de la violencia ha sido objeto de estudio de diversos investigadores de distintas áreas del conocimiento como la sociología, psicología, salud, economía, entre otras. Por lo que las posturas para su análisis y comprensión han sido variadas.

El estudio de la violencia ha sido un tema multidisciplinario que ha involucrado a científicos de una variedad de disciplinas, como la sociología, la psicología y la medicina, entre otras. La variedad de perspectivas teóricas ha permitido una mejor comprensión de la violencia desde diferentes puntos de vista y ha permitido abordarla en su complejidad.

Como se mencionó anteriormente, la sociología ha creado teorías que estudian la violencia en relación con la estructura social, el poder, la desigualdad y los procesos de cambio social. Los sociólogos como Émile Durkheim, Max Weber y Georg Simmel han ayudado a entender cómo la violencia se origina y se manifiesta en ciertos contextos sociales.

Por otro lado, la psicología ha estudiado la violencia desde una perspectiva individual, familiar y comunitaria, examinando elementos como la agresividad, la impulsividad, el trauma, el abuso y la dinámica interpersonal. Los psicólogos han creado teorías sobre cómo surge la violencia, cómo afecta a la mente la exposición a ambientes violentos y cómo prevenir y tratar la violencia.

Además, la salud pública ha abordado la violencia como un problema de salud global, investigando sus causas, efectos y métodos para prevenirla. Esto incluye el estudio de la violencia interpersonal, de género, comunitaria y estructural, así como la promoción de políticas y programas para reducir la incidencia y el impacto de la violencia en la salud de las personas.

Desde la perspectiva social, el tema de la violencia no puede verse aislado o atribuirse a una conducta individual, sino a un comportamiento que se condiciona a partir de factores socio-económicos.

De acuerdo con Briceño-León (2012), la falta de recursos en los individuos empuja a que se involucren en actividades delictivas y adopten una conducta delincuente a causa de las carencias que experimenta en su vida y la desigualdad percibida en la sociedad, donde observa que otros poseen lo que a él le falta.

Desde esta óptica, dado que las personas de bajos recursos carecen de acceso a los bienes necesarios para cubrir sus necesidades fundamentales, generadas muchas veces por la misma estructura social, y al no hallar una vía legal y establecida para obtenerlos, optan por apropiárselos mediante el uso de la fuerza.

Ha habido diversas formas de explicar el delito a través de la estructura social, tal como lo cita Briceño-León (2012) algunos enfoques se centran en analizar microsocialmente las comunidades urbanas, donde se identifican las condiciones de pobreza en zonas violentas. Estas condiciones pueden incluir el desempleo, bajos ingresos familiares, un alto porcentaje de jóvenes que no asisten a la escuela o una tasa alta de divorcios en las parejas. Estos factores se comparan con aquellos de otras zonas que no presentan altos índices de delincuencia.

Por su parte el enfoque macro social sostiene que existe una forma de violencia estructural en la sociedad. Según esta perspectiva, las condiciones de pobreza en sí mismas pueden considerarse una forma de violencia que se ejerce sobre las personas. Estas condiciones podrían explicar los comportamientos violentos de aquellas personas que, como víctimas de la pobreza, terminan convirtiéndose en agresores.

Dentro de los factores sociales que han sido utilizados por varios autores para la explicación de la violencia es la *desigualdad*, medida principalmente a través del índice de Gini, la cual surge de la insatisfacción de las personas en relación con otras que tienen más recursos y oportunidades. (Acevedo, 2008; Valiente, 2017; Carcach et al., 2011; Briceño-León, 2012; OMS, 2002)

La perspectiva de la *institucionalidad* social, desde la visión de Stein y otros (como se citó en Martínez, 2019), debe trascender la gestión para que las políticas y programas sean capaces de adaptarse a las nuevas expectativas y demandas y avanzar en la protección de los derechos. En otras palabras, las políticas sociales pueden cambiar de manera lógica y ordenada a través de procedimientos e instancias reconocibles y predecibles, lo que les da continuidad y capacidad de respuesta ante nuevos desafíos y demandas sociales.

La institucionalidad social se refiere al conjunto de normas formales e informales, incluidas las prácticas y prácticas organizacionales, que se utilizan para procesar y priorizar los problemas sociales y, al mismo tiempo, enmarcar el contenido y la dinámica administrativa y política de las políticas sociales. La institucionalidad social es vista por los actores de la política social como un conjunto de incentivos que establece sus límites y posibilidades de negociación y acción, de acuerdo con Repetto (como se citó en Martínez, 2019).

Un elemento importante de la Institucionalidad, es la consideración sobre el *territorio*, en el momento de desarrollar los atributos de la institucionalidad social. Esto está relacionado con la descentralización y las particularidades que conlleva la territorialidad para la política social (Martínez, 2019).

De acuerdo con Briceño-León (2012) los estudios que buscan comprender la violencia y el crimen a través de la interacción social de los individuos con el orden simbólico y las normas sociales que regulan su convivencia, ofrecen una perspectiva diferente sobre la violencia vista desde la institucionalidad. Es pues la institucionalidad, una variable a tomar en cuenta en el análisis de la violencia y sobre todo a la violencia homicida.

La perspectiva de la institucionalidad plantea una visión diferente en la “interpretación de los homicidios en América Latina y para ello se apoya en evidencias recientes”, afirma Briceño-León (2012). Es una alternativa a lo que se puede denominar la visión tradicional del análisis de la violencia.

Esta perspectiva de la institucionalidad afirma lo siguiente:

No es la pobreza ni la desigualdad lo que nos permite entender las variaciones de los homicidios, sino la institucionalidad de esas sociedades: el fortalecimiento o debilitamiento del orden normativo formal e informal que sirve de contención a los comportamientos violentos. (Briceño-León, 2012, p. 3160)

La adquisición de comportamientos delictivos se da a través de la interacción con otras personas y la elección entre obedecer o violar las normas y leyes. Para mantener la armonía en una sociedad, es necesario establecer un conjunto de reglas y regulaciones que regulen las relaciones sociales, asignen responsabilidades y establezcan lo que se considera comportamientos aceptables e inaceptables. De esta manera, se fomenta ciertos comportamientos y se criminalizan otros, lo que permite una relación social más fluida y predecible. (Briceño-León, 2012)

Sobre este punto Briceño-León (2012) continúa acotando que cada sociedad necesita un sistema de normas y leyes para controlar las interacciones sociales, asignar roles y determinar qué comportamientos son aceptables o condenables. Esto implica reconocer y promover ciertos comportamientos y castigar a otros para facilitar la convivencia y garantizar la previsibilidad del comportamiento de las personas.

A medida que la sociedad se vuelve más compleja debido a la especialización en la división del trabajo, se requieren mecanismos más rigurosos para mantener el orden social. Según los estudios de Durkheim, cuando estas normas y reglas no se transmiten o aprenden adecuadamente en una sociedad, pierde en su efectividad y se produce un estado de anomia. En este estado, las reglas sociales ya no guían el comportamiento y esto conduce a la aparición de comportamientos delictivos. (Briceño-León, 2012)

Merton retoma la perspectiva de Durkheim, pero la interpreta de manera distinta. En lugar de considerarla como un momento específico de la sociedad, la ve como una tensión que existe en todas las sociedades como resultado de la relación entre las metas que se consideran deseables para los individuos y la eficacia o legitimidad de los medios que se proporcionan para alcanzarlas.

Esta tensión en la sociedad se debe a la importancia que se le da a los objetivos y a las normas que regulan los medios para alcanzarlos. La respuesta de la sociedad ante aquellos que utilizan medios prohibidos para alcanzar los fines es crucial para mantener la norma en la sociedad. Si la sociedad no toma medidas contra aquellos que violan las normas, estas pierden

su valor y sentido, lo que lleva a la aplicación del castigo como un mecanismo de reciprocidad negativa para restablecer el equilibrio social.

El castigo es una forma negativa de expresar la norma social y ha existido en todas las sociedades como un medio para infligir dolor a aquellos que han causado dolor a otros. Al hacerlo, se restaura el valor universal de la norma social que ha sido violada.

Las personas evalúan la relación entre la posibilidad de alcanzar sus metas a través de medios proscritos y los costos que podrían resultar de dicha acción, incluyendo la probabilidad de que la sociedad responda con un castigo.

La relación entre estos dos factores es crucial y parece que la certeza de la aplicación del castigo, más que su severidad, refuerza el sistema normativo general. Esta perspectiva se ha aplicado también al estudio de los grupos sociales y su construcción de identidades en relación con las normas, y se ha utilizado para explicar la subcultura de la violencia y la creación de una cultura ciudadana.

2.1.1 Definición de violencia

La violencia, en su esencia, es un fenómeno complejo que puede manifestarse de múltiples maneras y que ha sido objeto de estudio y análisis por parte de diversos autores a lo largo del tiempo.

Una definición que tiene bastante aceptación es la que brinda la Organización Mundial de la Salud (OMS, 2002) en la que hace referencia a:

El uso deliberado de la fuerza física o el poder ya sea en grado de amenaza o efectivo, contra uno mismo, otra persona o de un grupo o comunidad, que cause o tenga muchas probabilidades de causar lesiones, muerte, daños psicológicos, trastornos del desarrollo o privaciones. (p. 3)

Esta definición es bien general y abarca los diferentes tipos de violencia: interpersonal, amenazas e intimidaciones, muerte, lesiones, psicológica, privaciones y deficiencias del desarrollo del bienestar de los individuos.

Según Anceschi (2009), la violencia es un concepto subjetivo de definición compleja, ya que puede adquirir diversos tipos de acepciones según el punto de vista desde el que se analice.

Calabrese (1997) indica que:

La violencia y la agresión son dos caras de la misma moneda que tradicionalmente ha sido aceptada como mecanismo de control por los individuos que han ostentado el papel hegemónico dentro del grupo social que de uno u otro modo se han visto justificados y, por lo tanto, legitimados en el ejercicio de esa violencia y de ese poder arbitrario. (p. 112)

Según Vidal (2008), la violencia es “la violación de la integridad de la persona”, la cual “suele ejercerse cuando interviene la fuerza física o la amenaza de su uso, pero también cuando se actúa en una secuencia que causa indefensión en el otro”, por lo que este autor la considera un proceso en el que participamos todos y no como un simple acto cuyo fin es la afirmación del “dominio” a través del cual busca el “control” de la presencia y las condiciones del estar, así como hacer del otro un medio considerándolo como propio y operando siempre sobre el “estar” del sujeto. (pp. 17-20)

De acuerdo con González et al. (2007):

Los principales factores de riesgo para que los adolescentes cometan actos violentos, incluido el homicidio, están dados por condiciones de violencia intrafamiliar grave; ser hombre; tener antecedentes graves de trastornos de la conducta, hiperactividad, grados altos de irritabilidad o intentos de suicidio; pertenecer a pandillas; permanecer en la calle la mayor parte del tiempo; consumir alcohol y SPA; usar armas; así como tener creencias que apoyan la resolución de problemas mediante la violencia, baja escolaridad y condiciones de pobreza importantes. (p. 92)

De acuerdo con lo anterior no hay una única definición de violencia, sin embargo, se podría decir que la violencia debe entenderse como una desviación del comportamiento socialmente aceptado que repercute negativamente la cohesión social de las estructuras de la sociedad y por tanto el tejido social.

La construcción del concepto de violencia, así como su explicación han sido abordadas desde diferentes enfoques entre ellos: económico, salud pública, ecológico, psicológico, sociológico, político y social; algunos de estos relacionados entre sí para proporcionar una comprensión más amplia de la violencia. Por tal razón, resulta que la construcción conceptual de la violencia lleva hablar de violencias y no solo de violencia, en la que convergen diversos elementos tanto en lo individual, como del entorno en el que la persona se desenvuelve.

2.1.2 Tipos de violencia

La *violencia interpersonal* de acuerdo con la OMS (2014), es la que ocurre entre miembros de una misma familia, en la pareja, entre amigos, conocidos y desconocidos, y que incluye el maltrato infantil, la violencia juvenil (incluyendo la asociada con pandillas), la violencia contra la mujer (por ejemplo, violencia de la pareja y la violencia sexual) y el maltrato de las personas mayores.

La *violencia juvenil*, se define como los “actos violentos propiciados por niños, adolescentes y jóvenes” (Interpeace-POLJUVE, 2009, p. 8). Por su parte, la Organización Panamericana de la Salud (OPS, 2003) señala que “la violencia juvenil es una de las formas de violencia más visibles en la sociedad y en casi todos los países del mundo” (p. 27).

Continuando con las definiciones de la OPS (2003), expresa que la *violencia colectiva* se define como:

El uso de la violencia como instrumento por parte de personas que se identifican a sí mismas como miembros de un grupo - ya sea transitorio o con una identidad más permanente - contra otro grupo o conjunto de individuos, para lograr objetivos políticos, económicos o sociales. (p. 235)

Otro tipo de violencia se menciona en el siguiente postulado que “la violencia cotidiana en las ciudades, es la causa de la mayoría de los homicidios y lesiones no fatales del país” (Duque y Klevens, 2000, p. 188).

Otro tipo de violencia es la denominada *violencia estructural*. Citando las palabras de Friedrich Engels escritas en 1844 (como se citó en La Parra y Tortosa, 2003), como conclusión a un informe sobre la situación de los barrios obreros de Londres:

Cuando una persona inflige un daño físico a otra, produciéndole la muerte, el acto es denominado homicidio sin premeditación; cuando el agresor conoce de antemano que el daño será fatal, llamamos a su acto, asesinato. Pero cuando la sociedad sitúa cientos de proletarios en una posición en la que de forma inevitable se encontrarán con una muerte prematura e inevitable (...), cuando priva a miles de personas de la satisfacción de las necesidades vitales, situándolos en condiciones en las que no es posible vivir — obligándolos, a través de la fuerza de la ley, a permanecer en esas condiciones hasta que la muerte sea la consecuencia inevitable—, la sociedad sabe que esos miles de víctimas perecerán y aun así permite que esas condiciones se mantengan, este acto es un asesinato

con tanta rotundidad como lo es el acto individual; asesinato disfrazado e intencionado contra el que nadie puede defenderse por sí mismo (...) porque nadie ve al asesino, porque la muerte de la víctima parece natural en tanto que el delito es más por omisión que por comisión. Pero asesinato, al fin y al cabo. (pp. 57-72)

Lo anterior es una forma de violencia directa, la cual es visible ante cualquier observador, con un agresor y una víctima que pueden ser identificados, pero también existe otra forma de violencia, que no es tan directa, más difícil de visualizar, en la que se produce un daño en la satisfacción de las necesidades humanas básicas es decir en la supervivencia, el bienestar, la identidad o libertad, como resultado de los procesos de estratificación social.

Esta forma de violencia invisible constituye a la violencia estructural, la cual para Galtung (1998) “es la suma total de todos los choques incrustados en las estructuras sociales y mundiales, y cementados, solidificados, de tal forma que los resultados injustos, desiguales, son casi inmutables” (p. 16).

La violencia estructural, teniendo en cuenta, además de los aportes de Galtung, el concepto desarrollado por La Parra y Tortosa (2003), quienes la definen como la “situación (...) que compromete las oportunidades de bienestar, identidad y/o libertad de las personas, de donde se originan condiciones de injusticia social, desigualdad, inequidad, pobreza o exclusión social” (p. 61), son factores que generan la inestabilidad, y facilitan el distanciamiento entre población y la institucionalidad.

Dentro de las explicaciones de la violencia estructural, La Parra y Tortosa (2003) citan a Coady, quien propone una alternativa de cambiar el término violencia y considerar los efectos de inestabilidad de las necesidades humanas como injusticia social, la desigualdad, inequidad, pobreza o exclusión social. Y plantean la siguiente definición sobre violencia estructural:

La idea de violencia estructural implica una ampliación semántica de la palabra violencia, cuyo objetivo es mostrar que su amenaza está presente de manera institucional incluso cuando no hay violencia en el sentido literal o “amplio”. La violencia estructural no involucra a actores que infligen daño mediante la fuerza, sino que es equivalente a injusticia social. Además de su potencialidad para llevar a confusión, el problema clave con el concepto es su dudosa sugestión de una variedad de problemas sociales que en apariencia son bastante diferentes son en realidad la misma cosa y tendrían que ser abordados de una única manera. (La Parra y Tortosa, 2003, p. 61)

Dentro de la discusión entre el uso de un término u otro, que incorpore en su definición la injusticia social o las formas de desigualdad, para La Parra y Tortosa (2003) el término violencia estructural sigue siendo el que mejor explica el conocimiento sobre esta. De tal forma que lo justifica de la siguiente manera:

- a) Tanto la violencia estructural como la violencia directa constituyen motivos por los que un importante sector de la humanidad no puede satisfacer sus necesidades humanas básicas.
- b) El término violencia estructural es útil para indagar la relación existente entre las situaciones de violencia estructural y formas de violencia directa. En este sentido, por poner un ejemplo, la violencia represiva puede entenderse como un mecanismo necesario para mantener una determinada estructura de poder cuando existen altos niveles de violencia estructural, de igual modo la llamada conflictividad social, ya sea de forma organizada (movimientos políticos) o desestructurada (delincuencia) se relacionan a menudo con respuestas a incrementos en los niveles de violencia estructural (por ejemplo, la negación del acceso a determinados bienes colectivos).
- c) El término violencia estructural es útil para introducir los mecanismos de ejercicio del poder como causantes de procesos de privación de necesidades humanas básicas. En efecto, la injusticia social, la pobreza o la desigualdad, no son fruto únicamente de dinámicas producidas por las relaciones de tipo económico, sino que también pueden ser explicadas a partir de la opresión política utilizando mecanismos tan dispares como la discriminación institucional, legislación excluyente de ciertos colectivos o la política fiscal y de gasto público regresiva, por citar algunos. Al hablar de violencia nos situamos en el campo semántico del poder con mayor facilidad que cuando se usan términos como pobreza o desigualdad, que frecuentemente han sido abordados desde lo económico, campo éste que, por supuesto, no puede ser obviado.
- d) Como podría ocurrir cuando en lugar del término violencia se utiliza el término desigualdad social, ya que se abren (aunque no necesariamente) ventanas para la legitimización de una situación en la que un grupo de personas tiene insatisfechas sus necesidades básicas. El término violencia estructural sirve por tanto para recordar que la eficiencia se produce en cualquier caso a costa de una forma de reparto que es sistemáticamente desfavorable para algunas de las partes, que esto es conflictivo y que

existen motivos para pensar que la situación es impuesta por los ganadores y no es deseada por los perdedores.

2.1.3 Variables que explican la violencia

2.1.3.1 Pobreza.

No existe una única definición de la pobreza, existen diversas posturas y puntos de vista sobre cómo debe entenderse, medirse y qué aspectos se deben tomar en cuenta para considerarla. Entre estos aspectos se pueden mencionar: la falta o insuficiencia de ingreso para garantizar medios de vida adecuados, hambre, malnutrición, falta de acceso o acceso limitado a los servicios de salud, educación, saneamiento y seguridad social; desempleo, carencia de vivienda, condiciones de inseguridad, discriminación o exclusión social.

Si bien en la actualidad existen diversos métodos para medir y determinar la pobreza en el país: la pobreza medida por el indicador de las necesidades básicas insatisfechas (INBI) y la pobreza desde una visión multidimensional. Algunos de estos métodos de medición de la pobreza son recientes y no se cuenta con mayores registros históricos de ellos, por lo que resulta difícil su comparación.

Tradicionalmente el método que se ha utilizado en El Salvador para medir y determinar la pobreza ha sido el **Ingreso**. A partir de éste, la pobreza se determina por el acceso que una persona o una familia puede tener a una Canasta Básica Alimentaria (CBA) y una Canasta Básica Ampliada. Así la pobreza puede ser absoluta si no se cubre una CBA; y relativa si se cubre una CBA, pero no una Canasta Básica Ampliada², ambas pobrezas forman lo que se denomina la línea de pobreza y se expresa en porcentaje.

Es importante mencionar que, el cálculo de la CBA la realiza mensualmente la Dirección General de Estadísticas y Censos (DIGESTYC)³ basándose en las calorías mínimas necesarias que necesita una persona para vivir al día.

2.1.3.2 Desigualdad.

América Latina y el Caribe es conocida como la región más desigual del mundo, de acuerdo con la Comisión Económica para América Latina y el Caribe (CEPAL, 2019):

² Una Canasta Básica Ampliada es el doble del valor de una CBA.

³ Se puede consultar el siguiente enlace: <http://www.digestyc.gob.sv/index.php/servicios/en-linea/canasta-basica-alimentaria.html>

La desigualdad de ingresos es una de las expresiones más evidentes de esa desigualdad y es un obstáculo al desarrollo y a la garantía de los derechos y del bienestar de las personas, además de un factor que inhibe la innovación, el aumento de la productividad y el crecimiento económico. (p. 21)

Generalmente cuando se habla de Desigualdad, se hace referencia a la concentración del ingreso en una población. Esto quiere decir que entre mayor sea la desigualdad, mayor será la concentración de los ingresos en pocas personas y entre menor sea la desigualdad, menor será la concentración de los ingresos en pocas personas.

Para medir esta desigualdad en la distribución del ingreso se utiliza el coeficiente de Gini, el cual considera la distancia de la curva de distribución del ingreso respecto de una línea teórica de equidistribución de este. El índice toma valores en el rango de 0 a 1, donde 0 corresponde a la equidad absoluta y 1 a la inequidad absoluta.

2.1.3.3 Desempleo.

De acuerdo con la DIGESTYC, el desempleo se calcula a partir de aquellas personas que cuentan con la edad legal para trabajar que no tienen trabajo, pero que andan en busca de uno de manera constante.

El desempleo se mide en porcentaje a través de la tasa de desempleo la cual expresa la proporción de la Población Económica Activa (PEA) que el sistema económico, en un periodo determinado no logra absorber.

2.1.3.4 Producto Interno Bruto per cápita (PIB).

De acuerdo con el Banco Central de Reserva (BCR), se define como PIB per cápita a la estimación promedio de la distribución del producto interno bruto entre el total de la población, en un período dado. El PIB es una medida del flujo total de bienes y servicios producidos por la economía de un país, durante determinado período, generalmente un año.

El PIB es uno de los indicadores que se ha utilizado para medir la riqueza de los países y al dividirlo entre sus ciudadanos permite observar cómo se distribuye esa riqueza entre las personas. De acuerdo con el Banco Mundial (2021): “La riqueza, al igual que el PIB, representa el bienestar material, no el bienestar humano considerado más ampliamente” (p. 2), por tanto, este indicador se centra únicamente en el aspecto material.

2.1.3.5 Índice de Desarrollo Humano (IDH).

El concepto de Desarrollo Humano fue introducido por el Programa de las Naciones Unidas para el Desarrollo (PNUD) en el año de 1990. Se entiende por Desarrollo Humano:

En primer lugar, el proceso por el cual se amplían las capacidades y las oportunidades del ser humano; y en segundo lugar, el nivel o estado de realización y bienestar que alcanza la persona humana como resultado de una adecuada aplicación de sus capacidades y de un correcto aprovechamiento de sus oportunidades. (PNUD, 1997, p. 21)

Dentro de este enfoque de desarrollo, se utiliza el Índice de Desarrollo Humano (IDH) que mide el nivel de vida de las personas de un país utilizando los siguientes indicadores: la esperanza de vida al nacer (longevidad), combinación del índice de alfabetismo con el índice de matriculación combinada (alfabetismo), y el poder adquisitivo sobre la base del PIB per cápita (vida decente). (PNUD, 1997)

El valor del IDH se calcula en un rango de 0 a 1, en el que aquellos valores cercanos a 1 tienen un mayor Desarrollo Humano. De acuerdo con este rango, se ha realizado una clasificación de países de la siguiente manera:

- Desarrollo humano muy alto, países con niveles superiores al 0.80.
- Desarrollo humano alto, países con niveles entre 0.70 y 0.79.
- Desarrollo humano medio, países con niveles entre 0.55 y 0.69.
- Desarrollo humano bajo, países con niveles inferiores a 0.55.

2.1.3.6 Institucionalidad o Gobierno de La Ley (Rule of Law).

La institucionalidad se refiere al estado de derecho o el gobierno de la ley, la cual se mide a través del índice sintético del **Rule of Law (Gobierno de la Ley)**, que construyó el Worldwide Governance Indicators (WGI), que evalúa el sistema normativo y de cumplimiento de la ley.

Este indicador recoge las percepciones que los diferentes agentes tienen en cuanto a la confianza y cumplimiento de las reglas de la sociedad, y en particular la calidad del cumplimiento de los contratos, los derechos de propiedad, la policía y los tribunales, así como la probabilidad de delitos y violencia. Su medición se hace a través de percentiles, el cual indica la ubicación de un país en relación con los demás, donde 0 corresponde al rango más bajo y 100 al rango más alto.

2.1.3.7 Violencia expresada en homicidios

La violencia se puede definir como el uso intencional de la fuerza física o el poder real o como amenaza contra uno mismo, una persona, grupo o comunidad que tiene como resultado la probabilidad de daño psicológico, lesiones, la muerte, privación o mal desarrollo.

De todos los actos violentos, el de mayor significación es el homicidio. Allí la víctima pierde la vida y por lo tanto es un acto final; se le atribuye la máxima significación, pues no hay retorno ni arreglo, se arrebató la existencia, se anula al otro.

“Tradicionalmente los análisis locales e internacionales de la violencia en El Salvador han tenido como indicador central los homicidios” (FESPAD, 2021, p. 25).

La medición de la tasa de homicidios se expresa a través del número de muertes por violencia por cada 100 mil habitantes y se expresa en porcentaje.

2.2 Fundamentos teóricos estadísticos

2.2.1 Regresión Lineal Simple

2.2.1.1 Análisis univariado.

Para el primer acercamiento a la exploración de los datos de las distintas variables consideradas en la investigación, se utilizará el análisis descriptivo, el cual define diversas medidas capaces de resumir la información recolectada a un pequeño número de valores. Estas medidas proporcionan una idea general de cómo se encuentran la distribución de los datos (Gorgas et al., 2011).

El análisis univariado el cual consiste en analizar de manera independiente cada una de las variables, examinando su comportamiento, en relación a sus medidas de tendencia central, de posición, de dispersión, asimetría y curtosis.

2.2.1.2 Análisis bivariado.

Dentro del estudio de las variables estadísticas bidimensionales o análisis bivariado, se analiza la existencia de relaciones o dependencias de una variable dependiente en relación con una o más variables independientes.

De acuerdo con Gorgas et al. (2011) la relación entre dos variables puede ser funcional o aleatoria, tal como lo describen a continuación:

La relación entre las dos variables podrá ser de dos tipos: funcional, cuando exista una relación matemática exacta que ligue ambas variables (ej. el radio y el área de un círculo), o aleatoria, cuando, aunque no exista entre las variables una relación exacta, se puede observar (aunque no siempre es el caso) una cierta tendencia entre los comportamientos de ambas (ej. el peso y la altura de un individuo). (p. 185)

2.2.1.3 Regresión lineal simple.

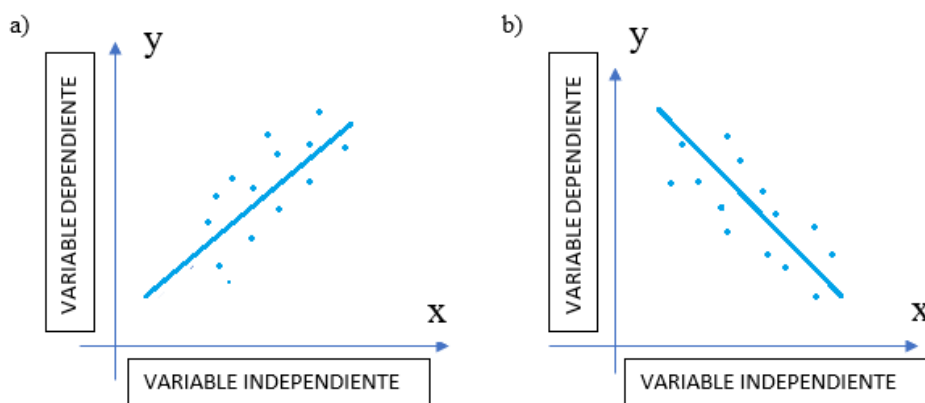
El análisis de regresión lineal simple se puede poner de manifiesto en la relación de dependencia sencilla, de tipo lineal, que se da entre dos variables cuantitativas, donde una de ellas es considerada como dependiente y la otra como independiente.

Parte de este análisis concierne también al estudio de la varianza, el objetivo consiste en alcanzar la mayor precisión en el conocimiento de la variable dependiente, es decir, se busca reducir la variabilidad entre variables, esto equivale a afirmar que una variable explica a otra variable, y también la predice.

El primer paso para el estudio de la relación entre las variables consiste en la construcción y observación de un diagrama de dispersión.

2.2.1.4 Diagrama de dispersión.

A partir de la consideración de dos variables, una dependiente y una independiente, la relación entre ambas se puede representar a lo largo del eje horizontal y la otra a lo largo del eje vertical, es frecuente que a la primera variable se le denomine x y a la otra y . El conjunto de puntos en dos dimensiones, se llama diagrama de dispersión. En la Figura 1, se muestra una representación de diagrama de dispersión.

Figura 1*Diagrama de Dispersión Lineal*

Nota. En la Figura 1 se presentan dos gráficas: a), muestra una tendencia lineal directa o positiva, es decir, cuando crece la variable independiente también crece la variable dependiente. La gráfica b), muestra los datos con una tendencia lineal inversa o negativa, es decir cuando la variable independiente crece la variable dependiente decrece y viceversa. Adaptada de “Estadística elemental (Lo esencial) (2ª Edición)”, R. Johnson, 1999, Ediciones Paraninfo.

En el diagrama de dispersión de acuerdo con Gorgas et al. (2011) se debe identificar el tipo de función (línea recta) que mejor se ajuste a la dispersión de todos puntos, una función que ajuste los valores observados con los valores que serán, por lo tanto, se conoce como línea de regresión a la representación gráfica de la función que se ajusta a la nube de puntos del diagrama de dispersión.

La regresión lineal de y sobre x viene dada por la siguiente fórmula o modelo:

$$y = a + bx \quad (1)$$

Esta recta es llamada también, Modelo Probabilístico Lineal Simple que considera el predecir el valor de la variable dependiente, basada en el valor de la variable independiente x , donde, a y b , son parámetros a estimar, gráficamente a , es la ordenada de la recta en el origen (es decir el valor de y para $x = 0$) y b es la pendiente de la recta estimada.

2.2.1.5 Coeficiente de correlación.

Además, este análisis de regresión está estrechamente relacionada con el cálculo de una medida de la relación entre variables cuantitativas determinado por el coeficiente de correlación de Pearson.

El concepto de correlación es una de las ideas básicas del análisis estadístico de datos a la hora de establecer relaciones entre dos variables cuantitativas, mide la fuerza o la intensidad de la relación entre dos variables, así como la dirección de esta, siempre y cuando esta relación sea de tipo lineal.

De acuerdo con Gorgas et al. (2011) la interpretación del coeficiente de correlación es el siguiente:

- a) Si el coeficiente de correlación $r = 0$, indica la ausencia de correlación (independencia entre las variables).
- b) Si el coeficiente de correlación $r = 1$, indica una dependencia funcional entre ambas variables y una correlación positiva, directa, perfecta.
- c) Si el coeficiente de correlación $r = -1$, indica una dependencia funcional entre ambas variables y una correlación negativa, indirecta, perfecta.
- d) Si el coeficiente de correlación $0 < r < 1$, indica que la correlación será positiva, pero no perfecta.
- e) Si el coeficiente de correlación $-1 < r < 0$, indica que la correlación es negativa, pero no perfecta.

En caso se obtenga un coeficiente de correlación bajo no significa que no pueda existir correlación entre las variables. De lo único que informa, es de que la correlación no es lineal (no se ajusta a una recta), pero es posible que pueda existir una buena correlación de otro tipo, de igual manera un coeficiente de correlación alto no significa que exista una dependencia directa entre las variables. Es decir, no se puede extraer una conclusión de causa y efecto basándose únicamente en el coeficiente de correlación (Gorgas et al., 2011, p. 195).

2.2.1.6 Coeficiente de determinación.

En general se puede utilizar el cuadrado del coeficiente de correlación R^2 , que es llamado también coeficiente de determinación, como una medida de la bondad de ajuste para la ecuación

de regresión estimada, es decir mide la relación entre las observaciones empíricas y el modelo estimado (Anderson et al., 2001).

El coeficiente de determinación expresa la porción de la variación total en los valores de la variable “y” que puede ser considerada o explicada por una relación lineal con los valores de la variable “x”, si $R^2 = 0.36$, significa que el 36% de la variación total de los valores de “y” en la muestra están explicados por la relación lineal con valores de “x” (Canavos, 1987, p. 395).

2.2.1.7 Análisis de la varianza para regresión lineal.

Este importante método de análisis de la varianza para la regresión, permite saber que tan bueno es el ajuste entre la regresión y los datos de la muestra. El procedimiento de análisis de varianza comienza por considerar, la variación total en el experimento, que es medida por la cantidad llamada **Suma Total de Cuadrados (STC)**.

“La **STC** se divide en la Suma Total de Cuadrados de la Regresión (**STCR**) o la variación explicada y la Suma de Cuadrados del Error (**SCE**) o la variación no explicada, conocida también como el Residual” (Mendenhall et al., 2017, p. 386).

Su cálculo se realiza a partir de la siguiente fórmula:

$$STC = STCR + SCE \quad (2)$$

La variación total consiste en la suma de los cuadrados de las diferencias entre la variable que se observa, la variable en estudio o encontrada en los registros (y) y el valor promedio de la variable respuesta (\bar{y}):

$$STC = \sum (y - \bar{y})^2 \quad (3)$$

La variación explicada, consiste en la suma de los cuadrados de las diferencias entre la variable respuesta que reporta el modelo o lo estimado (y') y el valor promedio de la variable respuesta (\bar{y}):

$$STCR = \sum (y' - \bar{y})^2 \quad (4)$$

La variación no explicada, consiste en la suma de los cuadrados de las diferencias entre la variable que se observa, la variable en estudio o encontrada en los registros (y) y la variable

respuesta que reporta el modelo o lo estimado (y'), se usa para medir la variación agrupada dentro de las k muestras:

$$SCE = \sum (y - y')^2 \quad (5)$$

De esta manera:

$$STC = STCR + SCE$$

$$\sum (y - \bar{y})^2 = \sum (y' - \bar{y})^2 + \sum (y - y')^2 \quad (6)$$

Por lo tanto, es necesario calcular solo dos de las tres sumas de cuadrados: STC, STCR y SCE y la tercera se puede hallar por sustracción. Cada una de las fuentes de variación: STC, STCR y SCE, cuando es dividida entre sus apropiados **grados de libertad**, da una estimación de la variación en el experimento. Como STC involucra n observaciones cuadradas sus grados de libertad son:

$$gl = n - 1$$

De igual manera la STCR comprende k observaciones cuadradas y sus grados de libertad son:

$$gl = k - 1$$

La SCE posee los siguientes grados de libertad:

$$gl = n - k$$

Se debe observar que:

$$gl = gl (STCR) + gl(SCE) \quad (7)$$

Estas fuentes de variación y sus respectivos grados de libertad gl se combinan para formar los cuadráticos medios. La variación total en el experimento se muestra en la siguiente tabla de análisis de varianza o por sus siglas ANOVA.

Tabla 1

Fórmulas para el Análisis de Varianza (ANOVA)

Fuente de variación	Suma de Cuadrados (SC)	Grados de Libertad (gl)	Varianza Cuadráticos Medios (CM)	Relación entre Varianzas (F_0)
Regresión	$STCR$ (VE)	$k - 1$	$CM_{reg} = \frac{STCR}{k - 1}$	$F_0 = \frac{CM_{reg}}{CM_e}$
Residual	SCE (VNE)	$n - k$	$CM_e = \frac{SCE}{n - k}$	
Total	STC (VT)	$n - 1$		

Nota. La Tabla 1 muestra el proceso y las fórmulas para el análisis de la varianza (ANOVA). Adaptada de “Probabilidad y estadística”, por G. Canavos, (1987), McGraw Hill.

2.2.1.8 Distribución F.

Para determinar si la correlación es significativa se utiliza la Distribución Fisher, la cual permite hacer un contraste para una prueba de hipótesis y verificar la existencia de correlación o una relación significativa entre dos factores.

La prueba consiste en comparar la variación explicada y la no explicada para contrastar la existencia de correlación y explicación.

Las hipótesis a considerar son:

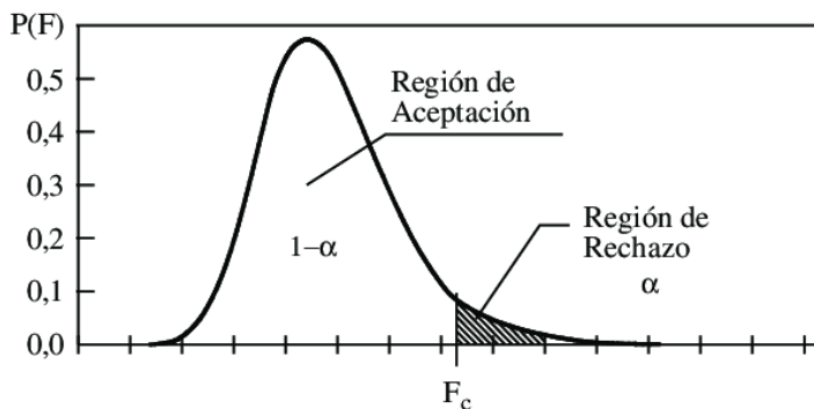
H_0 : No hay relación significativa entre x e y

H_1 : Hay relación significativa entre x e y

Suponiendo que x e y se encuentran normalmente distribuidos, α es el nivel de significancia que representa el riesgo de cometer el error del tipo I, $1 - \alpha$ es el nivel de confianza es decir la probabilidad de aceptar la hipótesis nula cuando es verdadera y α es la probabilidad de rechazar la hipótesis nula cuando es verdadera. Ver Figura 2.

Figura 2

Distribución F y la región de rechazo de la hipótesis



Nota. La Figura 2 muestra la gráfica correspondiente a la distribución F y la región de aceptación y rechazo de la hipótesis nula. De “Probabilidad y estadística”, por G. Canavos, 1987, McGraw Hill.

La distribución F se basa en los siguientes grados de libertad:

$$gl_1 = k - 1 \quad (8)$$

$$gl_2 = (b - 1)(k - 1) \quad (9)$$

El estadístico de prueba, para distribución F mediante el coeficiente de determinación es:

$$F_{cal} = \frac{r^2(n - 2)}{1 - r^2} \quad (10)$$

Donde n representa el número de datos que se tienen y el valor de F se incrementa conforme r^2 aumenta.

Cuan mayor sea la cantidad de variación de y que tenga explicación mediante la función de regresión ajustada, más probabilidad hay de que la prueba F considere a la regresión como significativa.

Se rechaza H_0 , si $F_{calculado} > F_{critico}$, donde $F_{critico}$ se encuentra en la cola superior de la distribución F . Por lo tanto, habrá suficiente evidencia para concluir que el modelo de regresión explica parte de la variación.

2.2.2 Regresión Lineal Múltiple

Los modelos de regresión múltiple pueden emplearse para predecir el valor de la variable dependiente o respuesta Y , o para evaluar la influencia que tienen los parámetros sobre ella.

Los modelos lineales múltiples siguen la siguiente ecuación:

$$Y_k = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + e_i \quad (11)$$

Donde $\beta_0, \beta_1, \beta_2, \dots, \beta_k$ son parámetros fijos, X_1, X_2, \dots, X_k se miden sin error y e , es una variable aleatoria que normalmente se distribuye alrededor de cero (media de e) y tiene una variación de V_e . Los valores específicos de e se obtienen mediante muestreo aleatorio simple e independiente, para una distribución norma $N(0, V_e)$ (Makridakis et al., 1978, p. 254).

Siendo β_0 la ordenada en el origen, el valor de la variable dependiente Y cuando todos los parámetros predictores son cero, además β_i , es el efecto promedio que tiene el incremento en una unidad de la variable predictora X_i sobre la variable dependiente Y , manteniéndose constantes el resto de variables y e_i es el residuo o error, la diferencia entre el valor observado y el estimado por el modelo.

Las condiciones necesarias y suficientes para la regresión múltiple se explican en los siguientes apartados.

2.2.2.1 Normalidad.

En los modelos lineales múltiples los residuos deben distribuirse de forma normal con media cero, para comprobarlo se recurre a histogramas o a test de hipótesis de normalidad.

Una prueba para determinar si los datos provienen de una población no normal, es el test de Shapiro-Wilk (Levin et al., 2014; Robbins, 2015), según Novales (2010), este test se emplea para contrastar normalidad cuando el tamaño de la muestra es menor a 50 observaciones y en muestras grandes es equivalente al test de Kolmogórov-Smirnov.

En este supuesto, la hipótesis nula, afirma que no existe correlación lineal si el p-valor es menor a alpha (nivel de significancia). En el supuesto de normalidad, se necesita

H_0 : Los datos provienen de una distribución normal.

H_1 : Los datos no provienen de una distribución normal.

2.2.2.2 Linealidad.

El aporte del psicólogo Charles Edward Spearman (Londres, 1863-1945), quien desarrolló el coeficiente de correlación ordinal que lleva su nombre, el cual permite correlacionar dos variables por rangos, en lugar de medir el rendimiento separado en cada una de ellas (Martínez et al., 2009).

El Coeficiente de correlación por jerarquías de Spearman (Rho de Spearman), es una medida de asociación lineal que utiliza los rangos, números de orden, de cada grupo de datos y compara dichos rangos, además es posible determinar también la dependencia o independencia de dos variables aleatorias (Elorza y Medina, 1999).

El Coeficiente de correlación por jerarquías de Spearman (Rho de Spearman) es recomendable utilizarlo cuando los datos presentan distribuciones no normales (Santander y Ruiz, 2004, pp. 1-30). Los rangos en los que puede variar el resultado del coeficiente que va desde -1 hasta 1, como puede verse en la Tabla 2.

Este coeficiente se calcula a partir de la siguiente fórmula:

$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} = \rho \quad (12)$$

Siendo:

n : la cantidad de observaciones

x_i : el rango de observaciones i con respecto a una variable

y_i : el rango de observaciones i con respecto a la segunda variable

$d_i = x_i - y_i$: es la diferencia entre los rangos de x e y (Anderson & Sweeney, 1999).

Tabla 2

Rangos en los que puede variar el coeficiente de correlación de Spearman

Rango	Relación
(-) 0.91 a (-) 1.00	Correlación negativa perfecta
(-) 0.76 a (-) 0.90	Correlación negativa muy fuerte
(-) 0.51 a (-) 0.75	Correlación negativa considerable
(-) 0.11 a (-) 0.50	Correlación negativa media
(-) 0.01 a (-) 0.10	Correlación negativa débil
0.00	No existe correlación
(+) 0.01 a (+) 0.10	Correlación positiva débil
(+) 0.11 a (+) 0.50	Correlación positiva media
(+) 0.51 a (+) 0.75	Correlación positiva considerable
(+) 0.76 a (+) 0.90	Correlación positiva fuerte
(+) 0.91 a (+) 1.00	Correlación positiva perfecta

Nota. La Tabla 2 muestra los rangos en los que puede variar el resultado del coeficiente de correlación por jerarquías de Spearman (Rho de Spearman). De “Metodología de la investigación”, por R. Hernández-Sampieri, 1998, McGraw-Hill.

En el supuesto de linealidad, se necesita

$$H_0: \rho = 0 \text{ (No existe correlación lineal)}$$

$$H_1: \rho \neq 0 \text{ (Existe correlación lineal)}$$

Donde ρ representa al coeficiente de correlación de Spearman. Para la construcción del valor crítico se requiere de:

n : cantidad de observaciones

gl : $(n - 2)$ grados de libertad

α : nivel de significancia (0.05)

Estos valores son necesarios para seleccionar en la tabla el valor crítico (t_c^*) y compáralo con el estadístico de prueba (t_c)

$$t_c = \frac{\rho}{\sqrt{\frac{1 - \rho^2}{n - 2}}} \quad (13)$$

Por lo tanto, no existe correlación lineal si el valor crítico (t_c^*) es mayor que el estadístico de prueba (t_c).

2.2.2.3 Independencia.

Los valores de cada observación deben ser independientes de los otros, en este supuesto la hipótesis nula indica que las variables son linealmente independientes si el Coeficiente de correlación por jerarquías de Spearman (Rho de Spearman) es mayor que el valor crítico, con una significancia de 0.05 (Elorza y Medina, 1999).

Para el supuesto de Independencia se plantea el siguiente sistema de hipótesis:

H_0 : Las variables son independientes

H_1 : Las variables no son independientes

2.2.2.4 Homocedasticidad.

La homocedasticidad es una propiedad deseable de los errores de un modelo de regresión, esto es, la varianza de los residuos se mantiene constante, es decir si la varianza es constante, se distribuyen de forma aleatoria manteniendo una misma dispersión y sin ningún patrón específico.

Esto significa que los errores de predicción o residuos e_i , no están relacionados con los valores predichos Y_i .

Para el caso de dos medias se utiliza tradicionalmente la prueba de t. Sin embargo, el análisis de varianza es una prueba más general que permite contrastar dos o muchas muestras (Sokal, 2002).

Para el supuesto de Homocedasticidad se tiene el siguiente sistema de hipótesis:

H_0 : La varianza de los grupos es igual

H_1 : La varianza de los grupos no es igual

2.2.2.5 Colinealidad.

En los modelos lineales múltiples las variables predictoras deben ser independientes, no debe haber colinealidad entre ellas. La colinealidad ocurre cuando una variable predictora está linealmente relacionada con una u otras variables predictoras del modelo, como consecuencia de la colinealidad no se puede identificar de forma precisa el efecto individual que tiene cada una de las variables colineales sobre la variable respuesta, lo que se traduce en un

incremento de la varianza de los coeficientes de regresión estimados hasta el punto que resulta prácticamente imposible establecer su significancia estadística, esto afecta negativamente a la precisión de la estimación de los parámetros.

Una primera aproximación de la multicolinealidad pasa cuando se observa la matriz de correlaciones, si se observa valores por encima de 0.85, entonces es probable que dichas variables estén aportando información redundante, otro método consiste en calcular las correlaciones múltiples al cuadrado entre cada variable y el resto, si se encuentra un valor por encima de 0.90 entonces hay sospecha de la presencia de la multicolinealidad en este caso multivariada (Kline, 1998).

La evaluación del modelo en conjunto, al igual que ocurre en los modelos lineales simples, el coeficiente de determinación r^2 , mide la bondad de ajuste del modelo, es decir permite conocer que tan bueno es el modelo para predecir el valor de las observaciones.

En los modelos lineales múltiples cuantas más variables predictoras incluya el modelo, mayor es el valor de r^2 , dado que cada variable o parámetro predictor explica una parte de la variabilidad observada en Y .

El valor de $r_{ajustado}^2$, depende del número de variables predictoras utilizadas y del tamaño de la muestra, es decir del número de grados de libertad, permite encontrar el mejor modelo aquel que consiga explicar mejor la variabilidad de Y con el menor números de variables predictoras.

El método para evaluar la bondad de ajuste esta dado por:

$$r_{ajustado}^2 = 1 - \left(\frac{SSE}{SST} \right) \frac{n-1}{n-k-1} = r^2 - (1-r^2) \frac{n-1}{n-k-1} = 1 - \frac{\frac{SSE}{df_e}}{\frac{SST}{df_t}} \quad (14)$$

Donde:

SSE es la variabilidad explicada por el modelo

SST es la variabilidad total de Y

n es el tamaño de la muestra

k el número de variables predictoras introducidas en el modelo

Para conocer la variabilidad que explica cada una de las variables predictoras del modelo se recurre a un ANOVA, ya que es el método que se encarga de analizar la varianza.

Tal y como ocurre en los modelos lineales simples o en los estudios de correlación, por muy alta que sea la bondad de ajuste, si el test F no resulta significativo no se puede aceptar el modelo como válido puesto que no es capaz de explicar la varianza observada mejor de lo esperado por azar.

2.2.2.6 Construcción de los modelos.

El análisis de regresión lineal múltiple es una técnica estadística utilizada para estudiar la relación entre variables en una amplia variedad de situaciones y predecir fenómenos diversos. La función de regresión más simple es la lineal, en donde cada variable participa de forma aditiva y constante para todo el fenómeno observado (Fernández, et al., 2006).

Un modelo estadístico es una expresión simbólica en forma de igualdad o ecuación que se emplea para indicar los diferentes factores que afectan o interviene en la variable de respuesta.

La construcción de un modelo indica, explicar todas las interrelaciones posibles e importantes implicadas en una situación, en tal caso es importante un proceso que permita seleccionar entre un modelo u otro el que mejor se ajuste a los datos.

En la selección de las variables predictoras se aplica el método paso a paso, el cual emplea criterios matemáticos para determinar qué variables predictoras contribuyen significativamente al modelo.

Dentro de este método se utiliza la estrategia de dirección *forward*: en el cual el modelo inicial no contiene ningún predictor, solo el parámetro β_0 y a partir de este se generan todos los posibles modelos introduciendo una sola variable de entre las disponibles, aquella variable que mejore en mayor medida el modelo se selecciona y continúa agregando variables predictoras restantes al modelo. Este proceso se repite hasta llegar al punto en el que ninguna de las variables que quedan por incorporar mejore el modelo.

El método paso a paso requiere de algún criterio matemático para determinar si el modelo mejora o empeora con cada incorporación o extracción.

Para ello los diseños factoriales son los más utilizados en investigaciones en las que se analizan dos o más factores, es decir, con dos o más variables independientes, las cuales serán

analizadas al ser consideradas como variables que afectan al proceso, en este tipo de diseño se investiga todas las combinaciones posibles (Montgomery, 2006).

Para ello se utiliza la fórmula, donde K es el número de regresores:

$$2^k \quad (15)$$

“Este procedimiento requiere que el analista ajuste todas las ecuaciones de regresión, que tengan un regresor candidato, dos regresores candidato, etc. Esas ecuaciones se evalúan de acuerdo con algún criterio adecuado y se selecciona el "mejor" modelo de regresión” (Montgomery, 2006, p. 271).

En los casos en los que se tenga que omitir alguna combinación posible, la familia de diseños factoriales $2^k - 1$ consiste en k factores, todos con dos niveles de prueba, en la que se omite la combinación de factores en su nivel más bajo.

2.2.3 *Métodos de selección de modelos*

Se seleccionarán aquellos modelos que cumplan con los criterios que se muestran a continuación:

2.2.3.1 **Criterio de Información de Akaike.**

Según Akaike (1974, pp. 716-723), proporciona un método simple y objetivo que selecciona el modelo más adecuado para caracterizar los datos experimentales. Este criterio, que se enmarca en el campo de la teoría de la información, se define como:

$$AIC = -2 \log(L(\hat{\theta})) + 2K \quad (16)$$

Donde $\log(L(\hat{\theta}))$, es el logaritmo de la máxima verosimilitud que permite determinar los valores de los parámetros libres de un modelo estadístico (Aldrich, 1997, pp. 162-176).

Una interpretación de AIC, considera el primer término de esta ecuación como una medida de la calidad con la que el modelo se ajusta a los datos experimentales, mientras que el segundo sería una penalización que se incrementa con la complejidad del modelo.

A manera de ejemplo, supóngase que, a partir de un conjunto de datos experimentales, se calculan los valores AIC obtenidos para dos modelos diferentes. El menor valor de AIC indica

que o bien el modelo se ajusta mejor a los datos experimentales o que es menos complejo o una combinación de ambos factores. Por lo tanto, este criterio ofrece un valor objetivo que, de manera relativa, cuantifica simultáneamente la precisión y sencillez del modelo.

2.2.3.2 Criterio de Información Bayesiana (BIC).

El criterio de información Bayesiana (BIC) propuesto por Schwarz en 1978, es una medida de bondad de ajuste de un modelo estadístico, y es utilizado para la selección de modelos entre un conjunto finito de modelos, tiene una penalización más fuerte que el AIC para modelos sobre parametrizados y ajusta el logaritmo de verosimilitud restringido -2 por el número de parámetros multiplicado por el logaritmo del número de casos, además está estrechamente relacionado con el criterio de información de Akaike. El BIC es definido como:

$$BIC = k \ln(n) - 2 \ln(L) \quad (17)$$

Donde: k , es el número de parámetros del modelo y $\ln(L)$, es la función de log-verosimilitud para el modelo estadístico.

Por lo que, para dos modelos estimados, el modelo con el menor valor de BIC es preferido; un BIC bajo implica un número menor de variables explicativas, mejor ajuste, o ambos.

2.2.3.3 Estadístico Predictivo Conceptual (C_p) de Mallows.

Es el método que comúnmente es utilizado para realizar todas las regresiones posibles y comparar el resultado, seleccionando un subconjunto de un conjunto de k regresores candidatos en un problema de regresión lineal múltiple, para un modelo particular con p parámetros.

El estadístico predictivo conceptual C_p de Mallows, es una forma de evaluar el ajuste de un modelo estimado al real, mediante el error cuadrático medio, donde se compara la precisión y el sesgo del modelo completo con modelos que incluyen un subconjunto de variables predictoras:

$$C_p = \frac{SSE}{\sigma^2} - n + 2p \quad (18)$$

Se han hecho varias sugerencias sobre cómo se debe interpretar exactamente la estadística, pero el consenso general es que los valores de C_p más pequeños son mejores, ya que indican cantidades más pequeñas de error no explicado.

Generalmente, debe buscar modelos en los que el valor de C_p de Mallows sea pequeño y esté cercano al número de predictores en el modelo más la constante (p). Un valor de C_p pequeño indica que el modelo es relativamente preciso (tiene una varianza pequeña) para estimar los coeficientes de regresión verdaderos y pronosticar futuras respuestas.

Un valor del C_p de Mallows que se aproxima al número de predictores más la constante indica que el modelo relativamente no tiene sesgo en la estimación de los verdaderos coeficientes de regresión y el pronóstico de respuestas futuras. Mientras que aquellos modelos con valores de C_p de Mallows más grandes que p , se consideran con falta de ajuste y sesgo.

2.2.4 Modelos lineales generalizados: distribución de Poisson

2.2.4.1 Función logaritmo.

Las variables respuestas son tratadas a menudo como variables continuas, en consecuencia, “el efecto de un conjunto de variables explicativas sobre dichas variables se analiza mediante un Modelo de Regresión Lineal (MRL)” (Oliver et al., 2005, p. 155).

Los modelos lineales son interpretados con mayor facilidad, sin embargo, pueden tener limitaciones importantes en capacidad predictiva, cuando los parámetros con los que cuenta no son lineales, resultando una aproximación demasiado simple para describir las relaciones reales entre variables.

La regresión no lineal permite encontrar una relación entre la variable dependiente y un conjunto de variables independientes, esto es ante el incumplimiento de los supuestos de homocedasticidad o normalidad y la naturaleza de las relaciones entre variables.

Los modelos lineales generalizados son modelos en los que la variable respuesta sigue una distribución diferente a la distribución normal. Eso contrasta con los modelos de regresión lineal, en los que la variable de respuesta sigue una distribución normal. Esto se debe a que los modelos lineales generalizados tienen variables de respuesta que son categóricas, como Sí, No; o Grupo A, Grupo B y, por lo tanto, no van de $-\infty$ a $+\infty$, por tanto, la relación entre la respuesta y las variables predictoras o explicativas puede no ser lineal.

Los modelos lineales generalizados, presentan tres componentes: el componente sistemático, es el predictor lineal que expresa la combinación lineal de las variables explicativas y proporciona el valor predicho, componente aleatorio: es un componente que recoge la variabilidad de Y no explicada por el predictor lineal y sigue una distribución de Poisson y la función de enlace, la cual enlaza el componente sistemático con el valor esperado es la función logarítmica.

Una alternativa de los modelos lineales generalizados es el Modelo de Regresión de Poisson, según Cameron (1998), dado que una variable respuesta es de naturaleza discreta, se ajusta a la distribución de Poisson bajo los supuestos de que los sucesos son independientes entre sí y de que la probabilidad de un suceso es constante durante el intervalo fijado.

El Modelo de Regresión de Poisson (MRP) es un caso particular del Modelo Lineal Generalizado, con distribución de Poisson para el componente aleatorio y con función de enlace logarítmica (Gill, 2001).

En General los modelos de *Regresión de Poisson* se utilizan mejor para modelar eventos en los que se cuentan los resultados, de manera específica se trabaja con datos discretos, con valores enteros no negativos, como la cantidad de veces que ocurre un evento durante un período de tiempo, por ejemplo, el número de veces que una persona es arrestada en un año, el número de patentes que solicita firma, el número de homicidios registrados en un día o en un año.

Los modelos seleccionados para el componente aleatorio de los Modelos Lineales Generalizados corresponden generalmente a distribuciones que pertenecen a la familia exponencial de distribuciones. De acuerdo con Everitt (2002) una distribución de probabilidad pertenece a la familia exponencial si su función de distribución de probabilidad posee la forma siguiente:

$$f(x) = \exp\{a(\theta)b(\theta) + c(\theta) + d(\theta)\} \quad (19)$$

Donde (θ) es un parámetro conocido como el parámetro natural de la distribución y a, b, c, d son funciones conocidas. La familia exponencial incluye a las distribuciones normal, gamma, binomial y Poisson.

La distribución de Poisson puede ser descrita de la manera siguiente:

$$f(x) = \exp\left\{(\ln \lambda)x - \lambda - \sum_{i=1}^n \ln(x_i)\right\} \quad (20)$$

Donde:

$$a(\theta) = \ln \lambda, \quad b(\theta) = x, \quad c(\theta) = -\lambda \quad y \quad d(\theta) = \sum_{i=1}^n \ln(x_i)$$

La derivación de la función de enlace requiere trabajar con la función de verosimilitud asociada a una muestra. Se asume que la variable aleatoria escalar y_i , dado un vector de regresores, x_i , y un vector de parámetros, β , está distribuida con densidad $f(y_i/x_i \beta)$.

El criterio de máxima verosimilitud elige como estimador de β , a aquel valor que maximiza una función de la probabilidad conjunta de los valores de y en la muestra, y_1, y_2, \dots, y_n .

Dicha función, vista como una función de los parámetros, condicional en los datos, es llamada la función de verosimilitud, denotada como:

$$L(\beta) = \prod_{i=1}^n f\left(\frac{y_i}{x_i} \beta\right) \quad (21)$$

En el caso de una distribución de Poisson, la función de verosimilitud asociada es

$$L(\beta) = \prod_{i=1}^n \frac{e^{-\lambda} \lambda^{y_i}}{y_i!} \quad (22)$$

Tomando el logaritmo natural de la función de verosimilitud y reemplazando $\ln \lambda$ con $x_i \beta$ se obtiene la función de enlace deseada:

$$\zeta(\beta) = \ln(L(\beta)) = \sum_{i=1}^n \ln\left(f\left(\frac{y_i}{x_i} \beta\right)\right) \quad (23)$$

$$\zeta(\beta) = \sum_{i=1}^n y_i x_i \beta - \exp(x_i \beta) - \ln y_i \quad (24)$$

En este caso interesa el MLG definido por:

$$\ln(\lambda) = x_i \beta \quad (25)$$

Esto define el Modelo Lineal Generalizado con componente aleatorio Poisson y función de enlace log, el cual también es conocido como regresión de Poisson.

Estas k ecuaciones son no lineales en las β desconocidas, y no hay solución analítica para $\hat{\beta}$, sin embargo, se pueden utilizar métodos iterativos, normalmente métodos gradientes tal como Newton-Raphson, son necesarios para calcular $\hat{\beta}$.

El modelo de Poisson es restrictivo en el supuesto de que la media es igual a la variancia. En situaciones prácticas es frecuente encontrar que la variancia de los datos excede a la variancia dicha en el modelo de Poisson. Ese problema puede surgir debido a factores tales como:

1. La población expuesta al riesgo es pequeña.
2. El número de homicidios en una comunidad es pequeño.
3. Los factores incluidos en la parte sistemática del modelo pueden mostrar patrones específicos de variación para diferentes grupos de la población o tipos de homicidios (Heterogeneidad).
4. Los eventos no son independientes.

Una manera de lidiar con el problema de exceso de dispersión es usando un modelo de probabilidad que tome en cuenta la asimetría y picudez de la distribución de la variable de interés en los datos. Uno de tales modelos es el de la distribución Binomial Negativa.

En Modelos Lineales Generalizados, la variable respuesta y_i se modela mediante una función lineal de variables explicativas y algún término de error:

$$y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + e_i \quad (26)$$

La salida y_i es un valor que sigue la distribución de Poisson. Asume el logaritmo de los valores esperados (media) que pueden modelarse en forma lineal mediante algunos parámetros desconocidos.

Para transformar la relación no lineal en forma lineal, se utiliza una función de enlace que es el logaritmo de la regresión de Poisson.

Por esa razón, un modelo de regresión de Poisson también se denomina modelo log-lineal. La forma matemática general del modelo de regresión de Poisson es:

$$\ln(y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + e_i \quad (27)$$

Donde:

- y es la variable respuesta
- β_0, β_i son coeficientes numéricos, β_0 es la intersección
- X_i variables predictoras/explicativas

Los coeficientes se calculan utilizando métodos como la **Estimación de máxima verosimilitud** o la cuasi-verosimilitud máxima.

Se considera una ecuación con una variable predictora y una variable de respuesta:

$$\ln(y) = \beta_0 + \beta_1 X_1 \quad (28)$$

Esto es equivalente a:

$$y = e^{\beta_0 + \beta_1 X_1} = e^{\beta_0} \cdot e^{\beta_1 X_1} = e^{\beta_0} \cdot (e^{\beta_1})^X \quad (29)$$

Por ello, a este modelo se le llama a menudo modelo log-lineal.

Aunque el supuesto fundamental para la aplicación correcta del MRP es el de la equidispersión, esto es, igualdad de media y varianza en las distribuciones condicionadas, habitualmente la varianza es superior a la media, denominando un fenómeno de sobre dispersión (Long, 1997).

Capítulo 3. Fundamentación Metodológica

En el marco de esta investigación, se ha definido un proceso metodológico que se basa en un enfoque cuantitativo y método hipotético deductivo. Este enfoque se seleccionó con el objetivo de permitir una exploración exhaustiva y precisa de las variables en estudio.

Para llevar a cabo este proceso, se realizó una búsqueda de los datos y bases de datos de información que se encuentran en páginas web, entre estas la del Banco Mundial (BM) (Worldwide Governance Indicators), Comisión Económica para América Latina y el Caribe (CEPAL), Programa de las Naciones Unidas para el Desarrollo (PNUD) y la Dirección General de Estadísticas y Censos (DIGESTYC), en combinación se han utilizado la estadística descriptiva e inferencial, para identificar posibles relaciones entre las variables analizadas y la incidencia de la violencia homicida en El Salvador, lo que ha permitido obtener conclusiones significativas y respaldadas por evidencia empírica, contribuyendo así a una comprensión más completa del tema de investigación.

La naturaleza de los datos de las variables en estudio sugiere la aplicación de la estadística descriptiva, así como de la estadística inferencial para su análisis. La estadística descriptiva ha sido utilizada para examinar y resumir las características básicas de las variables, como medidas de tendencia central, dispersión y distribución de los datos.

Por otro lado, la estadística inferencial ha sido empleada para realizar inferencias y sacar conclusiones sobre la población con base a una muestra de datos, utilizando técnicas como pruebas de hipótesis, análisis de correlación, entre otras. Esta combinación de enfoques estadísticos permitirá una comprensión más detallada de las relaciones y patrones presentes en los datos de la investigación, hasta derivar en un modelo estadístico que explique de mejor manera la comprensión de los homicidios en El Salvador.

3.1 Enfoque y método de investigación

3.1.1 Enfoque

La presente investigación busca acercarse a la realidad a través de un enfoque *cuantitativo*, que permita la observación empírica de los datos numéricos de una serie de variables para determinar posibles relaciones entre ellas. “La investigación cuantitativa

frecuentemente cuantifica relaciones entre variables, la variable independiente o predictiva y la variable dependiente o resultado” (Ábalos y Rebollo, 2022, p. 27).

La medición numérica y el análisis estadístico son aspectos elementales en el propósito de esta investigación. Sobre este punto, Fernández et al. (2014) expresa lo siguiente del enfoque cuantitativo: “Utiliza la recolección de datos para probar hipótesis con base en la medición numérica y el análisis estadístico, con el fin establecer pautas de comportamiento y probar teorías” (p. 4).

3.1.2 Método

Sobre el método de investigación, Echeverría (1999) propone una clasificación que trata de dar cuenta del pluralismo metodológico existente, y menciona los siguientes:

Los métodos deductivos e inductivos, los métodos de análisis y síntesis, los métodos experimentales, el método axiomático, los métodos matemáticos, los métodos de la observación, de medición, de clasificación, métodos heurísticos, el método hipotético-deductivo, métodos computacionales, y el método basado en la comprensión de los fenómenos estudiados. (pp. 116-117)

Por otra parte, Beltrán (2000) propone una tipología que reduce a cinco las vías de acceso a la realidad social: “el método histórico, método comparativo, método crítico-racional, método cuantitativo y cualitativo” (pp. 15-50).

En esta lógica de análisis deductivo, Urbano y Yuni (2014) plantean que “En la lógica cuantitativa, de corte hipotético-deductivo, todo el proceso de investigación se orienta a la verificación de las hipótesis” (p. 105).

Este estudio, de acuerdo a la naturaleza de las variables, se ha realizado desde la lógica del método *hipotético-deductivo*, entendiendo de manera concreta, la forma de conocer la relación entre la tasa de homicidios con la variable institucionalidad, medida a través del indicador estado de derecho (the rule and law) publicado por el Banco Mundial (Worldwide Governance Indicators); así como también con las variables: pobreza monetaria, la distribución del ingreso (índice de Gini), el desempleo, la riqueza nacional determinada por el Producto Interno Bruto (PIB) y el Índice de Desarrollo Humano (IDH), para luego establecer las magnitudes en las que se relacionan cada una de ellas con la violencia homicida.

3.2 Tipo de Investigación

3.2.1 Alcance de la investigación

El alcance de la investigación es *Correlacional*, lo que implica la asociación de variables a través de un patrón específico para un grupo determinado de población. Para medir estas variables, cada una se ha cuantificado y analizada su relación entre sí, para medir el grado de correlación entre ellas.

Según Fernández et al. (2014):

La utilidad principal de los estudios correlacionales es saber cómo se puede comportar un concepto o una variable al conocer el comportamiento de otras variables vinculadas. Es decir, intentar predecir el valor aproximado que tendrá un grupo de individuos o casos en una variable, a partir del valor que poseen en las variables relacionadas. (p. 94)

De acuerdo con Supo (2014) quien denomina estos estudios como de nivel relacional, plantea que se deben tener al menos dos variables, lo que da lugar a la comparación entre grupos antes y después, la asociación, la correlación y las medidas correspondientes. Todos estos procedimientos involucran dos variables y, por lo tanto, se les denomina análisis estadístico bivariado, lo que constituye la característica principal del nivel investigativo relacional.

3.2.2 Diseño de la investigación

El diseño de la investigación es *longitudinal*, puesto que recoge datos en un periodo de 30 años (de 1990 a 2019) de un total de seis variables. Para Fernández et al. (2014) los estudios con diseño longitudinal “recolectan datos en diferentes momentos o periodos para hacer inferencias respecto al cambio, sus determinantes y consecuencias” (p. 159). Sobre este mismo punto coincide Ábalos y Rebollo (2022) al decir que el diseño de investigación longitudinal “recolecta datos a través del tiempo en puntos o períodos especificados, para hacer inferencias respecto al cambio, sus determinantes y consecuencias” (p. 157).

Además, la investigación es *no experimental*, puesto que no se manipulará variable alguna para determinar resultados o condiciones de otras variables.

Podría definirse como la investigación que se realiza sin manipular deliberadamente variables. Es decir, se trata de estudios en los que no se hace variar en forma intencional las variables independientes para ver su efecto sobre otras variables. Lo que se hace en la investigación no experimental es observar fenómenos tal como se dan en su contexto natural,

para analizarlos (Fernández et al., 2014). Por esta razón estos estudios también son llamados observacionales (Urbano y Yuni, 2014; Pérez-Tejada, 2008).

El estudio de las variables es *retrospectivo*, ya que se analizan datos ocurridos en el pasado. Sobre este particular, Ganchozo et al. (2018) menciona que los estudios retrospectivos “son aquellos en los que se indaga sobre hechos ocurridos en el pasado” (p. 87).

3.3 Población y Muestra

La muestra para la presente investigación es una muestra no probabilística, en la cual, de acuerdo con Fernández et al. (2014):

La elección de los elementos no depende de la probabilidad, sino de causas relacionadas con las características de la investigación o los propósitos del investigador. Aquí el procedimiento no es mecánico ni se basa en fórmulas de probabilidad, sino que depende del proceso de toma de decisiones de un investigador o de un grupo de investigadores y, desde luego, las muestras seleccionadas obedecen a otros criterios de investigación. (p. 176)

El muestreo que se utilizó fue intencional, a partir de la revisión de la literatura y de las variables y datos de los cuales otros estudios ya han trabajado y que pueden ser consultados. Sobre este tipo de muestreo Muñoz (2015) señala lo siguiente:

En estas muestras el investigador elige razonada e intencionalmente las categorías que él considera representativas del fenómeno a investigar; sin embargo, aun cuando la selección es razonada o intencionada puede no ser absolutamente representativa y por ello carecer de objetividad, pues depende de la subjetividad del investigador. (p. 171)

Para esta investigación las variables seleccionadas, para explicar la violencia homicida son: gobierno de la ley, la pobreza, la distribución del ingreso, el desempleo, la riqueza nacional y el índice de desarrollo humano y entre los años 1990 y 2019.

Para este estudio se consideró un período de 30 años, considerando el período entre 1990 y 2019, de los cuales se recopilaron los datos en forma de mediciones e indicadores, que están relacionadas con el estudio como: la pobreza monetaria (pobreza extrema), el índice de Gini, desempleo, el índice de desarrollo humano, el PIB per cápita (Valiente, 2019; Montalvo, 2017; Carcach et al., 2011; Acevedo, 2008) y el estado de derecho o the rule and law, esta última variable es publicada por el Banco Mundial (Worldwide Governance Indicators).

Por lo que la muestra considerada en el presente análisis parte de una muestra de 30 datos por cada variable observada en el periodo de 1990 y 2019, algunas variables de las cuales no se pudo obtener una mediación por diversas situaciones, se utilizó el método de medios móviles, que son un indicador retardado, con lo que su objeto no es predecir cual es el comportamiento futuro de la tendencia, sino indicar cual está siendo su evolución (Olasolo et al., 2008, p. 58).

3.4 Técnica e instrumento de recolección de datos

La técnica principal que se utilizó fue la revisión documental, la búsqueda de información se realizó a través de fuentes secundarias principalmente en línea, dentro de estas fuentes se consultó a la Encuesta de Hogares de Propósitos Múltiples (EHPM) y los indicadores socioeconómicos de la base de datos del Banco Mundial (The Data Bank, World Wide Governance Indicators), a la que se accedió por medio del url: <https://databank.worldbank.org/home> y se pudo encontrar la información necesitada de manera anual.

Por la característica del tipo de investigación, no se utilizó un instrumento de recolección de datos propiamente dicho, sin embargo, se procedió a la elaboración de una base de datos comprende la información anual desde 1990 hasta 2019. En esta base de datos se vaciaron los datos de la tasa de homicidios, el indicador Gobierno de la Ley, de la pobreza extrema, la desigualdad (índice de Gini), el desempleo, el PIB y el IDH, la cual se constituyó en la base de datos con la que se han realizado todos los cálculos estadísticos. (Ver anexo 1)

3.5 Procesamiento y análisis de los datos

Los datos fueron analizados aplicando métodos estadísticos a través de las variables de interés en la investigación, usando la herramienta computacional de Excel, y haciendo un truncamiento de 2 cifras decimales, se espera que el enfoque mediante las variables de una realidad social ayude a mejorar el pensamiento crítico y las habilidades para resolver problemas en situaciones en las que se les presenten nuevas variables.

El análisis comienza abordando la hipótesis estadística como una manera formal de escribir una predicción sobre una población, y especificando la hipótesis nulas y alternativas que se pueden probar utilizando datos de muestra, para este estudio son:

H_0 : Gobierno de la Ley no es una variable significativa para explicar la variación en la tasa de homicidios en El Salvador dentro de un modelo estadístico que también incluye otras variables socioeconómicas como la pobreza extrema, la desigualdad, el desempleo, el PIB per cápita y el Índice de Desarrollo Humano.

H_1 : Gobierno de la Ley es una variable significativa para explicar la variación en la tasa de homicidios en El Salvador dentro de un modelo estadístico que también incluye otras variables socioeconómicas como la pobreza extrema, la desigualdad, el desempleo, el PIB per cápita y el Índice de Desarrollo Humano.

La hipótesis nula siempre predice ningún efecto o relación entre las variables, la hipótesis alternativa establece la predicción de su investigación de un efecto o relación.

Para decidir si aceptar o rechazar la explicación propuesta para el fenómeno, e investigar una relación entre las variables bajo estudio se han realizado una serie de fases.

Fase I. Análisis descriptivo

El primer paso que se realizó fue realizar un análisis descriptivo de cada variable bajo estudio, la tasa de homicidios, Gobierno de La Ley, la pobreza extrema, la desigualdad (índice de Gini), el desempleo, la riqueza nacional (PIB per cápita) y el índice de desarrollo humano (IDH).

Fase II. Análisis bivariado

Una vez conocido el comportamiento de cada variable bajo estudio, se realizó un análisis bivariado en donde la variable dependiente tasa de homicidios se asoció con cada una de las variables restantes, Gobierno de La Ley, la pobreza extrema, la desigualdad (índice de Gini), el desempleo, la riqueza nacional (PIB per cápita) y el índice de desarrollo humano (IDH) para obtener la matriz de correlaciones o/y asociaciones.

Fase III. Regresión múltiple

Se inicio el proceso de regresión múltiple, con el objetivo de construir un modelo lineal en el que el valor de la variable dependiente se determina a partir de un conjunto de variables explicativas o predictoras, en esta fase en las variables explicativas se verifican si existe colinealidad, independencia, variabilidad constante de los residuos, es decir la homocedasticidad y la normalidad de los residuos, en otras palabras, la normalidad de ellas.

Fase IV. Construcción de modelos

Para la construcción de los modelos, se empleó el método de los diseños factoriales, aplicando $2^k - 1$, a la formula original se le restó 1 para eliminar los modelos con una variable (constante) y partir de modelos con dos variables.

El resultado fue de 63 modelos, en los cuales se determinó las variables regresoras, coeficientes del modelo, la suma de cuadrados residuales $SSRes(p)$, el Coeficiente de determinación múltiple R^2 , R^2 ajustado y el cuadrado medio de residuales $MSRes(p)$, este método de paso a paso, emplea criterios matemáticos para decidir qué variables explicativas contribuyen significativamente al modelo y en qué orden se introducen.

Entre los criterios matemáticos para determinar si el modelo mejora o empeora con cada combinación de variables explicativas, se tomaron en cuenta el Criterio de Información de Akaike (AIC) y el Criterio de Información Bayesiano (BIC), los cuales son dos criterios de uso frecuente para la selección de modelos. De la misma forma se aplicó el criterio de Cp Mallows. Para los modelos obtenidos, se seleccionó el menor valor de AIC, lo que indica que o bien el modelo se ajusta mejor a los datos experimentales o que es menos complejo, y en realidad es una combinación de ambas variables, este criterio ofrece una precisión y sencillez del modelo. También se eligió el menor valor de BIC, lo cual es preferible, dado que un BIC bajo implica un número menor de variables explicativas, y un mejor ajuste. Para aplicar el criterio Cp de Mallows, se seleccionó el valor más pequeño y cercano al número de variables predictores, indicando que el modelo es relativamente preciso.

Fase V. Aplicación del modelo de regresión de Poisson

Dado que variable homicidios presenta una característica discreta medida en un periodo de tiempo, el modelo de regresión es apropiado. El modelo de regresión de Poisson (MRP) es un caso particular del Modelo Lineal Generalizado, dado que una variable respuesta es de naturaleza discreta, con función de enlace logarítmica.

3.6 Operacionalización de las variables

Tabla 3

Operacionalización de las variables

Variable	Definición	Dimensión	Indicadores	Unidad de medida
Violencia homicida	Expresada por los asesinatos cometidos por la violencia social.	Social	Tasa de homicidios	Tasa por cada cien mil habitantes
Efectividad de las instituciones públicas	Evalúa el sistema normativo y de cumplimiento de la ley en el país.	Político-jurídico	Gobierno de la Ley (The Rule and Law)	Percentil en un rango de 0 a 100 Donde 0 corresponde al rango más bajo y 100 al rango más alto.
Pobreza	Personas o familias que se encuentran por debajo de la línea de pobreza.	Socioeconómica	Ingreso	Porcentaje de la población total
Desigualdad	Se refiere a la desigualdad en el ingreso entre las personas o familias en una sociedad.	Socioeconómica	Índice de Gini	Se calcula en una escala de 0 a 100 Donde 0 representa igualdad perfecta y 100 implica desigualdad perfecta.
Desempleo	Se refiere a la proporción de la fuerza laboral que no tiene trabajo pero que está disponible y buscando empleo.	Socioeconómica	Tasa de desempleo	Porcentaje
Riqueza nacional	Es el valor monetario de todos los bienes y servicios finales producidos en el país en un determinado periodo de tiempo, dividido entre la población.	Económica	PIB per cápita	Dólares estadounidenses
Desarrollo humano	Evalúa el progreso a largo plazo en tres dimensiones básicas del desarrollo humano: una vida larga y saludable, el acceso al conocimiento y un nivel de vida digno.	Socioeconómica	Índice de desarrollo Humano (IDH)	Escala de 0 a 1 Donde 0 es el más bajo desarrollo humano y 1 el más alto desarrollo humano

Nota. La Tabla 3 muestra las variables que se consideran en el estudio, así como también los indicadores con los cuales se midieron.

Capítulo 4. Resultados

En este capítulo los resultados son presentados siguiendo la secuencia de los pasos mostrados en la metodología, retomando que las series de datos utilizadas son las siete variables con 30 observaciones cada una.

En la primera sección se hace un análisis exploratorio, presentado algunos estadísticos descriptivos, posteriormente se realiza un análisis bivariado de las variables.

Después, se llevan a cabo pruebas estadísticas para verificar los supuestos fundamentales de la regresión lineal, que incluye la normalidad, la linealidad, la independencia y la homocedasticidad.

Posteriormente, se construyeron 63 modelos de regresión lineal múltiple utilizando la fórmula $2^k - 1$. Estos modelos fueron evaluados utilizando los criterios de Cp, AIC y BIC, de los cuales se seleccionaron 10 modelos.

Finalmente, se aplicó el análisis de regresión de Poisson a los 10 modelos seleccionados y se probó uno a uno cada modelo con los valores de cada variable para el periodo de 1990 a 2019, con el objetivo de observar cuál modelo se acercaba más a la tasa de homicidio de cada año. Como resultado de esta prueba se determinó que el modelo que mejor explica la variabilidad en la tasa de homicidios está compuesto por las variables: Desigualdad (Gini) y PIB per cápita.

4.1 Análisis exploratorio de los datos

4.1.1 Análisis univariado

En el periodo de estudio (1990 a 2019) la media de la tasa de **pobreza extrema** fue de 40.4%, con una variación en dicho periodo entre un mínimo de 22.8% para el año 2019 y un máximo de 59.7% para los años 1990 y 1991. Otro valor a destacar es el de la curtosis, que para la pobreza extrema resultó en -0.55. Esto indica que la distribución de los datos es menos apuntada que una distribución normal. Es decir, la distribución tiene menos valores extremadamente altos o bajos de lo esperado bajo una distribución normal, lo que sugiere una menor concentración de datos en los extremos de la distribución. En lo que respecta al coeficiente de asimetría el resultado fue de 0.58, lo que indica que la distribución de los datos está sesgada hacia la derecha, aunque no de manera extrema, sino moderada. El resto de cálculos estadísticos descriptivos pueden observarse en la Tabla 4.

El comportamiento de la variable pobreza a lo largo del período estudiado se muestra en la Figura 3 gráfica A. Esta representación muestra claramente que no sigue un patrón completamente lineal; en cambio, muestra una serie de fluctuaciones que están marcadas por períodos de aumento y disminución. En general, se observa que ha habido una reducción considerable del porcentaje de familias en situación de pobreza a lo largo del periodo estudiado.

De acuerdo con estos datos se observa que en la década de los 90 hubo una reducción considerable de la pobreza en 18.0 puntos porcentuales. Mientras que en la década de los 2000 hubo una reducción de 1.0 punto porcentual; y para la década de 2010 al 2019 hubo una reducción de la pobreza de 13.0 puntos porcentuales.

En términos generales y haciendo una comparación entre estas tres décadas consideradas para este estudio, se podría determinar que en la última década un buen número de las familias que se encontraban en situación de pobreza, han salido de esta situación y por tanto cuentan con mejores condiciones de acceso a la CBA que en las dos décadas pasadas.

Estos resultados coinciden con lo expuesto por la Comisión Económica para América Latina y el Caribe (CEPAL, 2019) en los que se observa que en la región de América Latina:

Se produjo entre 2002 y 2014 un importante proceso de reducción de la pobreza y la pobreza extrema, así como de diversos indicadores de la desigualdad social. Este proceso estuvo asociado no solo a un contexto económico más favorable, sino también a un contexto político en que la erradicación de la pobreza y la disminución de la desigualdad social, así como el objetivo de ampliar la inclusión social y de extender la protección social, ganaron un espacio inédito en la agenda pública de muchos países de América Latina y, en cierta medida, del conjunto de la región. (p. 16)

Si bien entre 2015 y 2018 países como Brasil y Venezuela vieron un aumento en sus tasas de pobreza, en Centroamérica y México, la pobreza se redujo del 45% en 2014 al 42% en 2018 (CEPAL, 2019).

La pobreza puede verse afectada por una serie de factores socioeconómicos y políticos. Por ejemplo, por cambios en las políticas gubernamentales, fluctuaciones en los mercados laborales, eventos económicos globales, desastres naturales y otros fenómenos.

Es importante destacar que el análisis de estas fluctuaciones no solo proporciona información sobre el estado actual de la pobreza, sino que también permite identificar tendencias

a largo plazo y evaluar la efectividad de las políticas y programas diseñados para combatir la pobreza.

Por otra parte, la media de la **desigualdad** medida a través del índice de GINI fue de 0.47, con una variación en dicho periodo entre un mínimo de 0.38 para el año 2017 y un máximo de 0.54 para los años 1990, 1991 y 1998. En el caso del valor de la curtosis, para la variable desigualdad fue de -1.22, el cual a diferencia de la variable pobreza extrema, muestra una dispersión relativamente amplia de los datos y una cola más ligera en comparación con una distribución normal. El coeficiente de asimetría para esta variable fue de -0.36, plantea que la distribución de los datos está ligeramente sesgada hacia la izquierda, aunque no de manera extrema. El resto de mediciones pueden verse en la Tabla 4.

En la Figura 3 gráfica B, se muestra el comportamiento de la variable de desigualdad, medida a través del índice GINI durante el período examinado. Como lo demuestra esta representación visual, la tendencia no es completamente lineal; en cambio, se observan leves fluctuaciones que incluyen períodos de crecimiento y declive. En general, se observa una disminución significativa en la disminución de la desigualdad a lo largo del periodo estudiado.

Al analizar los resultados del índice de Gini para El Salvador entre los años 1990 y 1999, el índice se redujo en 2 puntos porcentuales, significando esto una mejora en la distribución del ingreso en esta década. La siguiente década hubo una mayor reducción de Gini, bajando aproximadamente 5.7 puntos porcentuales. Ya para la década de 2010 al 2019 Gini se redujo en 4.7 puntos porcentuales, un punto menos que la década anterior, pero sin embargo con una clara tendencia a la reducción.

Para poder tener un comparativo de cómo se encuentra la desigualdad de El Salvador en relación con algunos países de América Latina para el 2018, se tiene que “Los valores más bajos del índice de Gini, cercanos o inferiores a 0,400, se registran en la Argentina, El Salvador y el Uruguay, mientras que en el Brasil y Colombia los valores son superiores a 0,520” (CEPAL, 2019, p. 21).

El periodo analizado en esta investigación para El Salvador muestra una mejoraría en la distribución del ingreso medida por el Índice de Gini. Se puede decir que de acuerdo con los resultados de este indicador la sociedad salvadoreña va siendo cada vez menos desigual, puesto que para el año de 1990 el índice de Gini fue de 0.54 y para el año 2019 fue de 0.38, evidenciando notablemente una reducción considerable.

Estos resultados pueden contrastarse con lo que presenta la CEPAL (2019), en los que se indica que:

Entre 2014 y 2018, en cuatro países los niveles de desigualdad se redujeron de manera significativa según el índice de Gini. En Bolivia (Estado Plurinacional de), **El Salvador** y el Paraguay se produjeron disminuciones acumuladas del 7% o más, en tanto que en Colombia se alcanzó un 2%. Por su parte, el Brasil fue el único país en que se registró un aumento del índice de Gini superior al 3%. (p. 22)

En cuanto al **desempleo**, se observó que el comportamiento estadístico-descriptivo de esta variable entre los años 1990-2019, presenta una media de 6.16%, con una variación en dicho periodo entre un mínimo de 3.69% para el año 2013 y un máximo de 9.94% para el año 1994. El valor de la curtosis fue de -0.75, lo que evidencia una dispersión considerable de los datos y una cola menos pronunciada en comparación con una distribución normal estándar. El coeficiente de asimetría para esta variable fue de 0.02, esto significa que la distribución de los datos tiene una concentración similar de datos en ambos lados de la media y que la cola de la distribución no está sesgada hacia la derecha ni hacia la izquierda de manera significativa. Los datos se muestran en la Tabla 4.

En la Figura 3 gráfica C, se observa el comportamiento del desempleo durante el período estudiado, el cual muestra una dinámica no lineal, sino más bien describe una línea con una serie de variaciones que la vuelven irregular. Estas variaciones se deben a una variedad de factores económicos, políticos y sociales que afectan las tasas de desempleo en diferentes momentos.

A pesar de estas variaciones, se observa una tendencia general hacia una disminución del desempleo durante el período analizado. La disminución en las tasas de desempleo podría deberse a una serie de políticas y medidas implementadas para fomentar el empleo, así como a cambios en la situación económica que han creado más posibilidades laborales.

Al revisar los resultados de la tasa de desempleo de El Salvador, desde el año 1990 hasta el 2019, se observa una notable reducción de 3.3 puntos porcentuales en 30 años.

Al observar detalladamente cada década contemplada en este estudio, de 1990 a 1999 hubo una reducción de la tasa de desempleo de 0.85 puntos porcentuales. En la siguiente década el comportamiento en el desempleo fue menos favorable, sobre todo por la crisis mundial generada en 2008, tal como lo plantea la CEPAL/OIT (2009) en su boletín Coyuntura Laboral en América Latina y El Caribe:

La crisis económica continuó golpeando a los países de América Latina y el Caribe durante el segundo trimestre de 2009. Los mercados internacionales se mantuvieron débiles, lo que afectó las exportaciones regionales de bienes y servicios, las remesas y la inversión extranjera directa continuaron cayendo, el crédito perdió dinamismo y la masa salarial se contrajo, sobre todo por la pérdida de empleo. (p. 1)

La crisis repercutió directamente en el aumento de la tasa de desempleo en 0.37 puntos porcentuales al final de la década. Para la década de 2010 a 2019, el desempleo mostró un comportamiento parecido a la década de 1991 a 1999, mostrando una reducción de 0.72 puntos porcentuales.

Al contrastar estos resultados obtenidos de los Indicadores del Desarrollo Mundial (IDM) con los resultados de la Encuesta de Hogares de Propósitos Múltiples (EHPM) de la DIGESTYC⁴, en esta se refleja que para el año 2005 la tasa de desempleo era de 7.2% y para el año 2009 una tasa de 7.3% mostrando un aumento en una décima porcentual, coincidiendo con el aumento en la tasa para el cierre de esta década. Mientras que en el 2010 fue de 7.1% y en 2019 fue de 6.3%, describiendo una reducción de 0.8 puntos porcentuales, parecido comportamiento a las tasas de los IDM.

Es importante señalar que la medición del **Producto Interno Bruto (PIB) per cápita**, uno de los indicadores examinados en este estudio, tiene el registro oficial más antiguo, remontándose a 1962.

Por otra parte, los resultados obtenidos mediante el análisis estadístico descriptivo del PIB per cápita arrojan que la media fue de USD 2505.87 entre 1990 y 2019. Desde 1990 hasta 2019, hubo una gran variación desde un mínimo de USD 914.13 hasta un máximo de USD 4167.73. En cuanto al valor de la curtosis de -1.13, lo cual indica que la distribución de los datos presenta una dispersión notable y una cola menos pronunciada en comparación con una distribución normal estándar. Esto sugiere que la concentración de datos alrededor de la media es menos aguda y que hay menos valores extremos en los extremos de la distribución en comparación con lo que se esperaría en una distribución normal estándar. El coeficiente de asimetría para esta variable fue de 0.09, este valor significa que hay una ligera concentración de

⁴ El sitio web de la DIGESTYC únicamente tiene las publicaciones de las EHPM a partir del año 2005. Por esta razón no se pudo comparar la década de 1990 a 1999.

datos hacia los valores más altos, con una cola que se extiende un poco más hacia los valores superiores en comparación con los valores inferiores. El resto de valores descriptivos se muestran en la Tabla 4.

En la Figura 3 gráfica D, se observa el comportamiento del PIB per cápita durante el período estudiado, el cual muestra una dinámica no lineal, sin embargo, si describe una trayectoria más cercana a una recta positiva. En el periodo observado de 1990 a 2019 se aprecia una tendencia ascendente del PIB per cápita en El Salvador.

Al realizar un análisis de la década 1990 a 1999 este indicador aumentó en USD 1,016.50 por persona. En la siguiente década 2000 a 2009 el PIB per cápita aumentó USD 856.94, teniéndose un menor crecimiento en relación con la década anterior. Se debe recordar que es en esta década, específicamente en 2008 que ocurre la crisis mundial, la cual afectó considerablemente las economías no sólo de El Salvador, sino de varios países. Para la década de 2010 a 2019 hubo un aumento de USD 1,184.50, dicho aumento superó las dos décadas anteriores.

En cuanto al **Índice de Desarrollo Humano (IDH)**, se observó el comportamiento estadístico-descriptivo entre los años 1990-2019, obteniéndose una media de 0.62, con una variación en dicho periodo entre un mínimo de 0.53 para el año 1991 y un máximo de 0.67 para los años 2010-2019. El valor de la curtosis de -0.70, sugiere que la distribución de los datos muestra una dispersión significativa y una cola menos pronunciada en comparación con una distribución normal estándar. En lo que respecta al coeficiente de asimetría de -0.79, indica que hay una concentración significativa de datos hacia los valores más bajos, con una dispersión más amplia hacia los valores más altos. Por lo que la distribución de los datos está sesgada hacia la izquierda de manera considerable. El resto de valores descriptivos se muestran en la Tabla 4.

En la Figura 3 gráfica E, se muestra el comportamiento del IDH de El Salvador, sin embargo, describe un comportamiento cercano a una recta con tendencia positiva creciente.

Al observar el comportamiento del IDH en El Salvador durante la década de 1990 a 1999, este aumentó de 0.53 a 0.60, es decir que hubo un incremento de 0.07 finalizando la década con un nivel de desarrollo medio. Para la década del 2000 a 2009 se mejoró este indicador en 0.04, aunque el aumento fue inferior a la década anterior, este indicador describió un compartimento positivo.

Para la década siguiente, 2010-2019, la situación del IDH ha sido menos favorable, puesto que apenas creció 0.005 en diez años, muy por debajo del crecimiento que el indicador mostró en las dos décadas analizadas anteriormente. Sin embargo, la clasificación se mantiene en un nivel de desarrollo medio.

Al comparar este indicador con otros países, puede verse una tendencia similar hacia el mejoramiento del nivel de vida de las personas. Por ejemplo, países como Perú, Ecuador y Colombia que pueden clasificarse con un desarrollo humano alto; mientras que República Dominicana, El Salvador, Bolivia, Paraguay, Honduras, Nicaragua y Guatemala se clasifican con un desarrollo humano medio (PNUD, 2013, pp. 157-158).

A nivel de las regiones del mundo puede observarse que América Latina ha tenido una mejoría con un IDH de 0.741, sólo a unos puntos por detrás de la región de Europa y Asia Central (con 0.771) y muy por encima de otras regiones en desarrollo como Asia Oriental y el Pacífico (0.683), Estados Árabes (0.652), Asia Meridional (0.558) y África Subsahariana (0.475) (PNUD, 2013-2014).

El indicador Institucionalidad, conocido también como **Gobierno de la Ley**, se empieza a medir en El Salvador a partir del año 1996 por el Banco Mundial.

Los resultados obtenidos del análisis estadístico descriptivo del indicador Gobierno de la Ley pueden observarse en la Tabla 4. La media fue de 28.3 puntos entre 1990 y 2019, el puntaje menor se observó en el año 2018 con 19.7 puntos, mientras que el puntaje mayor se observó en el año 2004 con un puntaje de 40.2. El cálculo de la curtosis de -0.77, sugiere que la distribución de los datos exhibe una dispersión considerable y una cola menos pronunciada en comparación con una distribución normal estándar. El coeficiente de asimetría para esta variable fue de 0.50, lo que significa que la distribución de los datos está sesgada hacia la derecha, pero no de manera extrema.

Por otra parte, el comportamiento gráfico del Gobierno de La Ley es no lineal, tal como puede verse en la Figura 3 gráfica F, se observan fluctuaciones en el periodo en estudio.

A partir de la década de los 90, el indicador inicia con 22.1 y termina en el año 1999 con 28.1, mostrando un crecimiento de 6 puntos percentiles.

Para la década del 2000 a 2009 el comportamiento del indicador es variado, entre subidas y bajadas, en este periodo (2004) se observa el mayor punto de elevación alcanzando su máximo

puntaje. Sin embargo, a partir del 2005 hasta el 2009 el comportamiento fue hacia la baja, terminando la década con 27.5 percentiles.

La década siguiente (2010 a 2019) se inicia con 24.2 percentiles (3.3 percentiles menos que el año anterior), pero en los siguientes años de 2011 a 2014 el indicador aumenta hasta llegar a 36.5 percentiles. Este mejoramiento en el indicador coincide con la denominada tregua entre pandillas apoyada por el Gobierno de El Salvador, la cual inició en marzo de 2012 y terminó en enero de 2015. Es a partir de este año que el indicador inicia un descenso, llegando a su nivel más bajo registrado en 23 años con 19.7 percentiles en 2018. Para 2019 el indicador inicia una leve recuperación de 3.9 percentiles.

En este primer análisis descriptivo de los datos, se observa que ninguna de las variables descritas presenta una distribución normal perfecta. Algunas de ellas, como la Pobreza e IDH, muestran valores de curtosis más cercanos a cero, lo que indica una distribución más cercana a la normalidad en términos de forma de campana. Sin embargo, al examinar el coeficiente de asimetría, se observa que las variables desempleo y PIB per cápita tienen valores más próximos a cero, lo que sugiere una simetría en la distribución de los datos.

Tabla 4*Estadísticos descriptivos de las variables en estudio en el periodo 1990 a 2019*

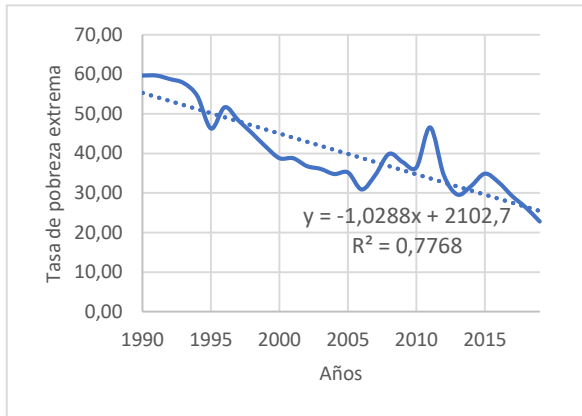
Estadísticos	Media	Error típico	Mediana	Moda	Desviación estándar	Varianza de la muestra	Curtosis	Coficiente de asimetría	Rango	Mínimo	Máximo	Suma	Cuenta
POBEXT	40.41	1.88	37.30	59.70	10.28	105.61	-0.55	0.58	36.90	22.80	59.70	1212.23	30.00
GINI	0.47	0.01	0.48	0.54	0.05	0.00	-1.22	-0.36	0.16	0.38	0.54	14.11	30.00
DESEM	6.16	0.30	6.49	7.53	1.63	2.64	-0.75	0.02	6.25	3.69	9.94	184.88	30.00
PIB per cápita	2505.87	178.84	2353.50	#N/D	979.55	959520.15	-1.13	0.09	3253.60	914.13	4167.73	75176.14	30.00
IDH	0.63	0.01	0.64	0.66	0.05	0.00	-0.70	-0.79	0.14	0.53	0.67	18.83	30.00
G&L	28.31	1.08	27.80	22.10	5.93	35.20	-0.77	0.50	20.50	19.70	40.20	849.40	30.00
H	76.85	5.55	64.64	#N/D	30.41	924.67	-0.31	0.90	105.72	36.00	141.72	2305.45	30.00

Nota. En la Tabla se encuentran los resultados de la estadística descriptiva de las variables Pobreza Extrema medida a través del Ingreso; el índice de GINI, que mide la desigualdad; el Desempleo; el Producto Interno Bruto per cápita (PIB per cápita); el Índice de Desarrollo Humano (IDH); el Gobierno de la Ley (G&L); y los Homicidios (H).

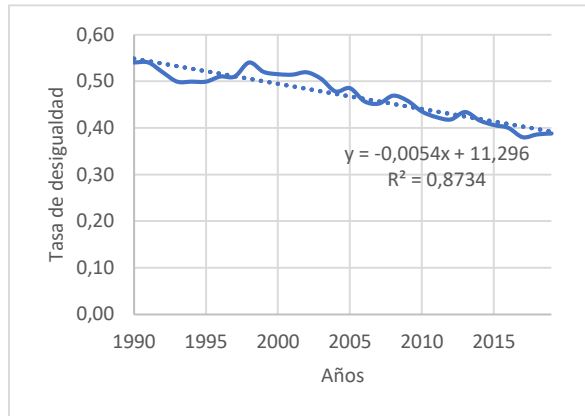
Figura 3

Representación gráfica del comportamiento de las variables en estudio en el periodo 1990 a 2019

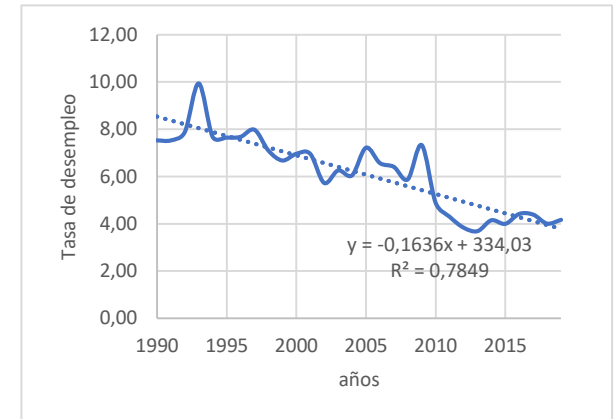
Gráfica A. Pobreza extrema



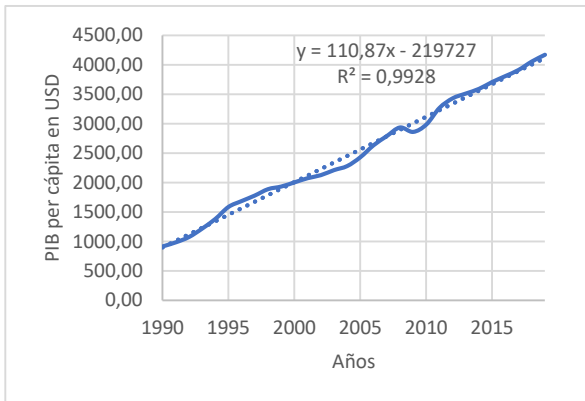
Gráfica B. Desigualdad (GINI)



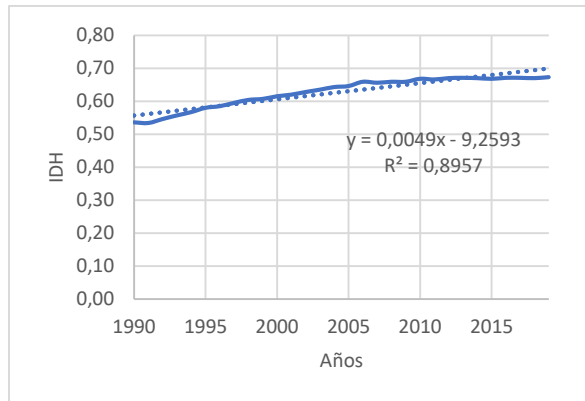
Gráfica C. Desempleo



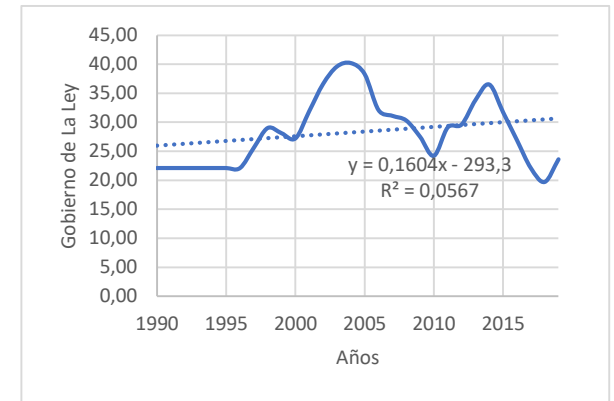
Gráfica D. Producto Interno Bruto (PIB) per cápita



Gráfica E. Índice de Desarrollo Humano (IDH)



Gráfica F. Gobierno de La Ley



Nota. Las figuras describen el comportamiento gráfico que presentan cada una de las variables consideradas en este estudio.

Por su parte la variable **Homicidios** en el periodo de estudio (1990 a 2019) presenta una media de la tasa de homicidios de 76.8%, con una variación en dicho periodo entre un mínimo de 36.0% para el año 2019 y un máximo de 141.7% para el año 1995. El valor de la curtosis de -0.31, se sugiere que la distribución de los datos presenta una dispersión notable y una cola menos acentuada en comparación con una distribución normal estándar. En cuanto al valor obtenido del coeficiente de asimetría de 0.89, es bastante cercano a uno, lo que indica un sesgo moderado hacia la derecha en la distribución de los datos, por lo que hay una concentración moderada de datos hacia los valores más altos, con una dispersión un poco más amplia hacia los valores más bajos. El resto de cálculos estadísticos descriptivos pueden observarse en la Tabla 5.

Tabla 5

Estadísticos descriptivos de la variable Homicidios en el periodo 1990 a 2019

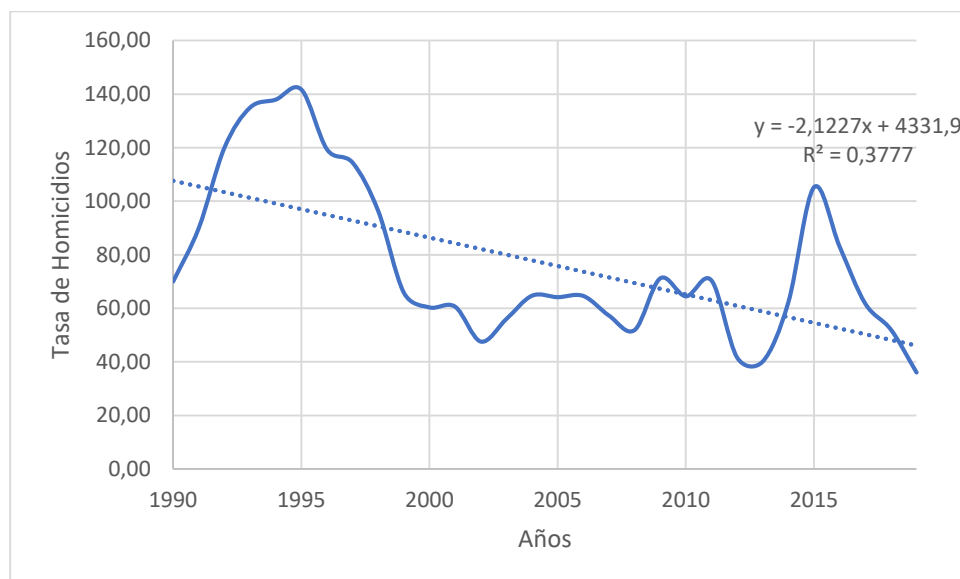
Estadístico	Resultados
Media	76.84
Error típico	5.55
Mediana	64.64
Moda	#N/A
Desviación estándar	30.41
Varianza de la muestra	924.66
Curtosis	-0.31
Coeficiente de asimetría	0.89
Rango	105.72
Mínimo	36
Máximo	141.72
Suma	2305.45
Cuenta	30.00

Nota. En la Tabla se encuentran los resultados de la estadística descriptiva de la variable Homicidios.

El análisis visual de los homicidios a través de la Figura 4 revela un patrón de comportamiento no lineal, evidenciando una serie de fluctuaciones marcadas por períodos de ascensos y descensos a lo largo del período estudiado.

Figura 4

Gráfica del comportamiento de la variable Homicidios en el periodo 1990 a 2019



Nota. La Figura muestra el comportamiento de la tasa de homicidios en El Salvador durante el periodo de 1990 al 2019.

De acuerdo con el periodo en estudio, en la primera década (1990 a 1999) se observan las tasas más elevadas, esto se debe a que el país aún se encontraba en guerra, la cual terminó hasta el año 1992.

En la década del 2000 al 2009, se mantienen tasas elevadas de homicidios, pero atribuidos a varios fenómenos sociales como el secuestro, droga, crimen organizado y la proliferación de pandillas, que cada vez aumentaría su presencia y criminalidad en todo el país.

En la siguiente década que va de 2010 al 2019, esta se caracteriza por el aumento del control territorial por parte de las pandillas, pero además por su incidencia política con los principales partidos políticos del país. Esta incidencia lleva a que en el año 2012 se establezca una tregua entre pandillas y por tanto se dé una considerable disminución de homicidios que se mantendría hasta finales del año 2013.

A partir de esa ruptura se presenta un aumento en las tasas de homicidios, lo que colocó a El Salvador con uno de los mayores índices de criminalidad de América Latina. Convirtiendo el año 2015 como el más violento, después de la firma de los Acuerdos de Paz, con un “registro histórico de 103 homicidios por cada 100 mil habitantes” (Comisión Interamericana de Derechos Humanos, 2021).

Esta variable tampoco presenta un comportamiento similar al de una distribución normal, ya que el valor de curtosis es mayor que cero, pero negativos, lo que indica una distribución sesgada hacia la izquierda. Además, el coeficiente de asimetría es bastante cercano a uno, lo que sugiere un sesgo pronunciado hacia la derecha. Estas discrepancias en los indicadores de la distribución refuerzan la idea de que la variable no sigue una distribución normal.

4.1.2 Análisis bivariado

4.1.2.1 Coeficiente de correlación.

La aplicación del coeficiente de correlación de Pearson con respecto a los homicidios, para el análisis bivariado permitió observar la magnitud en la que las variables se correlacionan linealmente, ya sea positiva o negativamente, de esta manera se hace el análisis de las variables predictoras con los homicidios.

Como puede observarse en la Tabla 6, dentro del conjunto de variables consideradas en el estudio, la variable que mayor correlación presenta es la **Pobreza** con una correlación positiva de 0.71, esto indica que en la medida que la pobreza aumenta, aumentan los homicidios.

La variable **Índice de Desarrollo Humano (IDH)** presenta la segunda correlación más alta del conjunto de variables, con una correlación negativa de -0.66, esto significa que en la medida que el IDH disminuye (lo que implica un menor desarrollo humano) los homicidios aumentan, obteniéndose una relación inversa entre el desarrollo humano y la incidencia de homicidios.

En tercera posición se encuentra la variable **Desempleo**, la cual presenta una correlación positiva de 0.64, lo que significa que en la medida que aumenta el desempleo, aumenta los homicidios. Por lo tanto, cuando hay un aumento en el desempleo, es probable que también haya un aumento en la incidencia de homicidios.

El **Producto Interno Bruto (PIB) per cápita** se ubica en la cuarta posición de las correlaciones, presenta una correlación negativa de - 0.58, lo que sugiere que en la medida que

disminuye el PIB per cápita, aumentan los homicidios. Consecuentemente, cuando la economía experimenta una disminución en el PIB per cápita, es probable que también haya un aumento en la incidencia de homicidios.

Tabla 6

Correlaciones de Pearson entre las variables explicativas y los homicidios

<i>Variables</i>	<i>H</i>	<i>GINI</i>	<i>DESEM</i>	<i>PIB</i>	<i>IDH</i>	<i>POBEXT</i>	<i>G&L</i>
H	1						
GINI	0.39739625	1					
DESEM	0.64143187	0.81595626	1				
PIB	-0.5869478	-0.9407508	-0.8920824	1			
IDH	-0.6686511	-0.8195936	-0.8164074	0.93308907	1		
POBEXT	0.71057001	0.72883412	0.75181755	-0.8672201	-0.9143574	1	
G&L	-0.4720497	-0.0023519	-0.2480959	0.19656375	0.44536089	-0.4395643	1

Nota. La Tabla muestra las correlaciones entre las variables en estudio, la variable que mayor correlación presenta con la variable homicidios es la pobreza.

Las variables que menos correlación presentan en el conjunto de variables analizadas son: **Gobierno de la Ley** e **Índice de GINI**.

La variable **Gobierno de la Ley** presenta una de las correlaciones más baja con relación a la tasa de homicidios, solo por encima del índice de GINI. Sin embargo, es importante destacar que existe una correlación negativa de -0.47, la cual no es nada despreciable dentro del análisis.

Una correlación de -0.47 indica una relación inversa moderada entre la variable Gobierno de la Ley y la tasa de homicidios. Esto significa que a medida que la calidad del Gobierno de la Ley mejora, es probable que la tasa de homicidios disminuya. Aunque la correlación no es tan fuerte como con las otras variables consideradas en el estudio, sigue siendo una relación importante a considerar en el análisis, ya que sugiere una relación importante entre el estado de la gobernanza y la incidencia de homicidios.

La variable **Índice de Gini** presenta una correlación positiva de 0.39; siendo una de las que menor correlación presenta en relación con la tasa de homicidios. Esto significa que a medida que la desigualdad social medida por el Índice de Gini aumenta, también tiende a aumentar la

incidencia de homicidios. Aunque la correlación no es tan fuerte como la de otras variables consideradas en el estudio, sigue siendo significativa y sugiere una asociación importante entre la desigualdad social y la incidencia de homicidios.

Dado que los resultados obtenidos muestran valores diferentes de cero, sugiere la presencia de alguna relación entre las variables estudiadas. La sola existencia de correlación no asegura que las variables sean adecuadas para modelos de regresión, por lo que se necesita ampliar el análisis a otras pruebas estadísticas como la validez de los supuestos de regresión lineal.

Las correlaciones entre las variables predictoras pueden verse que son altas, lo que supone colinealidad, lo que hace pensar que este supuesto no lo superaran estos modelos.

4.1.2.2 Coeficiente de determinación.

Se aplicó el coeficiente de determinación al conjunto de variables en análisis, para observar aquellas que mejor explican la incidencia de los homicidios. El resultado de este coeficiente se interpreta mejor cuanto más cercano esté a 1, lo que indica un mejor ajuste del modelo a los datos y una mayor capacidad del modelo para explicar la variabilidad en la variable dependiente. La Tabla 7 muestra los resultados obtenidos de este coeficiente.

Tabla 7

Coeficiente de determinación entre las variables en estudio

VARIABLES	R²
H vs GINI	0.15
H VS DESEM	0.41
H VS PIB	0.34
H VS IDH	0.44
H VS POBEX	0.50
H VS G&L	0.22

Nota. La Tabla muestra el coeficiente de determinación entre las variables en estudio, la variable que presenta mayor coeficiente de determinación es la variable pobreza.

El orden de los resultados del coeficiente de determinación, no varían en relación al orden de los resultados de las correlaciones de la Tabla 6, destacando en primer lugar a la variable pobreza con el R^2 más alto, lo que significa que explica aproximadamente el 50% de la variabilidad en la tasa de homicidios. Esto sugiere que la pobreza tiene una influencia significativa en la incidencia de homicidios, y que cambios en la tasa de pobreza están asociados con cambios sustanciales en la tasa de homicidios

Por su parte, la variable Índice de Desarrollo Humano (IDH) tiene un R^2 de 0.44, lo que indica que explica aproximadamente el 44% de la variabilidad en la tasa de homicidios. Aunque ligeramente menor que la pobreza, el IDH sigue siendo un predictor importante de la incidencia de homicidios.

La variable desempleo tiene un R^2 de 0.41, lo que significa que explica aproximadamente el 41% de la variabilidad en la tasa de homicidios. Esto sugiere que el desempleo también juega un papel significativo en la incidencia de homicidios.

En cuanto a la variable Producto Interno Bruto (PIB) tiene un R^2 de 0.34, lo que significa que explica aproximadamente el 34% de la variabilidad en la tasa de homicidios. Aunque aún importante, su capacidad explicativa es menor que las variables anteriores.

En lo que respecta a la variable Gobierno de Ley tiene un R^2 de 0.22, lo que indica que explica aproximadamente el 22% de la variabilidad en la tasa de homicidios. Aunque sigue siendo significativa, su influencia es menor que las variables anteriores.

La variable Índice de Gini tiene el R^2 más bajo, con solo el 15% de la variabilidad en la tasa de homicidios explicada por esta variable. Esto sugiere que el índice de desigualdad de ingresos tiene la menor incidencia en la tasa de homicidios entre las variables consideradas.

Con base a estos resultados, las variables con coeficientes de determinación más altos explican mejor la variación en la incidencia de homicidios. Estas variables han demostrado tener una mayor capacidad para predecir y explicar las fluctuaciones observadas en las tasas de homicidios, lo que las hace candidatas importantes para ser incluidas en modelos estadístico.

4.2 Regresión Lineal Múltiple

Un primer paso es realizar el análisis de regresión lineal múltiple a las variables contempladas para la presente investigación. Como técnica estadística implica, primero, la

comprobación del cumplimiento de los siguientes supuestos estadísticos tales como: normalidad, linealidad, independencia, homocedasticidad y multicolinealidad.

4.2.1 Normalidad de las series

Una prueba para determinar si los datos provienen de una población no normal es el test de Shapiro-Wilk (Levin et al., 2014; Robbins, 2015), según Novales (2010), este test se emplea para contrastar normalidad cuando el tamaño de la muestra es menor a 50 observaciones (en el caso de esta investigación es la que aplica por tratarse de una muestra de 30 elementos) y en muestras grandes es equivalente al test de Kolmogórov-Smirnov. Las estadísticas son presentadas en la Tabla 8.

Tabla 8

Resultados del Test de Shapiro -Wilk para analizar normalidad en la serie de datos

N°	VARIABLES	W-stat	p-valor	Alpha	Normal
1	GINI	0.92	0.03	0.05	No
2	DESEM	0.92	0.03	0.05	No
3	PIB	0.96	0.26	0.05	Sí
4	IDH	0.85	0.00	0.05	No
5	POBEXT	0.93	0.05	0.05	Sí
6	G&L	0.92	0.03	0.05	No
7	H	0.88	0.00	0.05	No

Nota. La Tabla 8, muestra los valores obtenidos del test Shapiro-Wilk evidenciado cuales variables tienen un comportamiento normal y cuáles no.

De acuerdo con los resultados observados en la Tabla 8, se puede ver que las variables GINI, Desempleo, IDH, Gobierno de la Ley y Homicidios, no presentan normalidad en los datos, ya que el p-valor es menor a alpha (nivel de significancia). Mientras que los datos del PIB y la Pobreza extrema, sí presentan normalidad, ya que el valor de la significancia (alpha) es menor que el p-valor.

Lo que evidencia que no todas las variables consideradas para esta investigación provienen de una distribución normal. Debido a que no se cumple este supuesto, para el siguiente supuesto se utilizó un test para datos no paramétricos.

4.2.2 Linealidad

El coeficiente de correlación de Spearman, es una medida de asociación lineal que utiliza los rangos, números de orden de cada grupo de datos y compara dichos rangos, además es posible determinar la dependencia o independencia de dos variables aleatorias, según Elorza y Medina (1999).

El coeficiente de correlación de Spearman es recomendable utilizarlo cuando los datos presentan distribuciones no normales, de acuerdo con autores como Santander y Ruiz (2004, pp. 1-30).

Según Hernández y Fernández (1998), la interpretación de los valores se ha expresado en escalas, siendo una de las más utilizadas la que se presenta en la Tabla 9.

Tabla 9

Resultado del test de Spearman para analizar correlación en las series de datos

Variable (y)	Variables (x)	Rango	CC-Spearman	Relación
H	GINI	(+) 0.11 a (+) 0.50	0.36	Correlación positiva media
	DESEM	(+) 0.51 a (+) 0.75	0.70	Correlación positiva considerable
	PIB	(-) 0.51 a (-) 0.75	-0.60	Correlación negativa considerable
	IDH	(-) 0.51 a (-) 0.75	-0.64	Correlación negativa considerable
	POB	(-) 0.51 a (-) 0.75	0.68	Correlación negativa considerable
	G&L	(+) 0.91 a (+) 1.00	0.95	Correlación positiva perfecta

Nota. En la Tabla 9 se observa que todas las variables en estudio tienen algún grado de correlación con la variable Homicidios (H), siendo la que mayor correlación positiva presenta, la variable Gobierno de la Ley (G&L) y la que menor correlación tiene GINI.

Si bien, de acuerdo a los resultados anteriores, la correlación más alta es la variable Gobierno de la Ley, seguida del Desempleo, la Pobreza Extrema y el IDH, se debe considerar que la dirección de las correlaciones no es en un único sentido.

En este supuesto de linealidad, la hipótesis nula indica que no existe correlación lineal si el valor crítico (t_c^*) es mayor que el estadístico de prueba (t_c) (13).

Los resultados del test y sus correspondientes valores críticos son presentados en la Tabla 10.

Tabla 10

Estadística de contraste del test de Spearman para analizar correlaciones lineales

Variable (y)	Variabes (x)	(t_c^*)	(t_c)	Resultado
H	GINI	± 2.0484	2.0386	*No existe correlación lineal
	DESEM		5.1291	Existe correlación lineal
	PIB		-3.9501	Existe correlación lineal
	IDH		-4.4133	Existe correlación lineal
	POB		4.9263	Existe correlación lineal
	G&L		15.6314	Existe correlación lineal

Nota. La Tabla 10 muestra la estadística de contraste del test de Spearman para analizar la linealidad.

Concluyendo que existe evidencia estadística para mostrar que existe una correlación lineal entre las variables bajo estudio, rechazando la hipótesis nula:

$$H_0: \rho = 0 \text{ (No existe correlación lineal)}$$

4.2.3 Independencia

Para el supuesto de independencia, la hipótesis nula indica que las variables son linealmente independientes si el Coeficiente de Spearman es mayor que el valor crítico, (t_c^*) (13) con una significancia de 0.05:

$$H_0: Rho > t_c^* \text{ (existe independencia)}$$

Los resultados de la prueba estadística para cada una de las variables en estudio se muestran en la Tabla 11.

Tabla 11

Estadísticas del test de Spearman para analizar independencia

Variable (y)	Variables (x)	(t_c^*)	CC-Spearman (Rho)
H	GINI	± 2.0484	0.3595
	DESEM		0.696
	PIB		-0.5982
	IDH		-0.6405
	POB		0.6814
	G&L		0.9472

Nota. La Tabla 11 muestra la estadística de contraste del test de Spearman para analizar la independencia.

Al comparar el valor crítico y el coeficiente de correlación de Spearman se evidencian los siguientes resultados:

1. Para la variable GINI combinada con la variable Homicidio, presentan independencia, por tanto, se acepta la hipótesis nula.
2. Mientras que el resto de variables no presentan independencia. Por tanto, se rechaza la hipótesis nula.

4.2.4 Homocedasticidad

Para el supuesto de homocedasticidad, se verificó si las distribuciones tienen la misma varianza, aplicado el análisis de varianza bivariado de un factor (ANOVA), en donde la hipótesis nula indica que las varianzas son iguales si el p-valor es menor que el nivel de significancia de 0.05:

$$H_0: p - \text{valor} < 0.05 \text{ (medias iguales)}$$

Los resultados de la prueba estadística pueden observarse en la Tabla 12 que se presenta a continuación.

Tabla 12*Resultado del análisis de la varianza de cada serie de datos*

Variable (y)	Variabes (x)	p-valor	Alpha	Resultado
H	GINI	7.63E-20	0.05	Las varianzas de cada grupo son diferentes.
	DESEM	2.41E-18		
	PIB	1.17E-19		
	IDH	8.37E-20		
	POB	6.66E-08		
	G&L	7.57E-12		

Nota. La Tabla 12 muestra los resultados estadísticos del análisis bivariado de las varianzas de la variable Homicidios con las demás variables en estudio.

Con base a los resultados de la Tabla 12, se evidencia que el valor p es menor que el nivel de significancia de 0,05. Por tanto, se rechaza la hipótesis nula, y se concluye que las varianzas son diferentes.

4.2.5 Multicolinealidad

Para analizar adecuadamente los resultados de un modelo de regresión lineal múltiple, es importante estudiar el grado de relación lineal existente entre las variables explicativas que lo componen.

La multicolinealidad es la relación de dependencia lineal fuerte entre más de dos variables explicativas en una regresión múltiple que incumple el supuesto de Gauss-Markov cuando es exacta.

La multicolinealidad es un problema que surge cuando las variables explicativas del modelo están altamente correlacionadas entre sí, el análisis de la matriz de correlaciones entre las variables exploratorias es una medida de diagnóstico de la multicolinealidad.

La existencia de correlaciones altas, mayores que 0.8 o 0.9, entre variables explicativas es también una señal de multicolinealidad alta. Los resultados de las correlaciones se muestran en la Tabla 13.

Tabla 13*Resultado del análisis de correlación Rho de Spearman*

Variables	H	GINI	DESEM	PIB	IDH	POBEXT	GL
H	1.00						
GINI	0.35	1.00					
DESEM	0.70	0.73	1.00				
PIB	-0.60	-0.90	-0.88	1.00			
IDH	-0.59	-0.89	-0.90	0.98	1.00		
POBEXT	0.68	0.77	0.80	-0.88	-0.85	1.00	
GL	-0.46	-0.06	-0.40	0.28	0.34	-0.40	1.00

Nota. La tabla presenta las correlaciones entre todas las variables en estudio, puede observarse que hay correlaciones altas entre las variables explicativas, por lo que la colinealidad está presente en este conjunto de variables.

Luego de haber aplicado diversos estadísticos para comprobar los supuestos, se concluye que:

Las únicas variables que describen una distribución normal son las variables PIB per cápita y Pobreza extrema. Por otra parte, al analizar las correlaciones del conjunto de variables en estudio, la variable desigualdad (GINI) presenta una correlación baja, por lo que no sería apropiada para considerarse en un modelo lineal. Los cálculos del supuesto de homocedasticidad arrojaron que, todas las variables presentan varianzas diferentes, por lo que no se cumple este supuesto para el conjunto de variables. Finalmente, en el caso de la multicolinealidad, se tiene que la correlación entre las variables independientes es alta, por lo que existe multicolinealidad entre las variables.

Los resultados obtenidos en este apartado, sugieren la consideración de modelos lineales generalizados. Sin embargo, para este caso particular, se optó por construir modelos de regresión lineal múltiple, aun cuando no se cumplieron todos los supuestos, con el objetivo de practicar la discriminación o eliminación de variables independientes como parte del proceso de construcción del modelo. Los modelos lineales generalizados se retoman en el apartado siguiente.

4.3. Modelo Lineal Múltiple

Luego de realizados la exploración de los datos y la verificación de los supuestos, se continua con la construcción de los modelos de regresión lineal múltiple. Para ello, se inicia con las diferentes combinaciones entre las variables regresoras, que en total fueron 63.

4.3.1 Selección de los modelos

Para la exploración de la combinación de las variables se utilizó la formula: $2^k - 1$, los resultados se presentan en la Tabla 14.

Este enfoque sistemático y exhaustivo en la construcción de combinaciones facilita la identificación de patrones, tendencias o relaciones que podrían no ser evidentes al examinar las variables de manera individual.

Esta combinación es una herramienta valiosa para explorar la complejidad de las relaciones entre las variables y contribuye significativamente a la comprensión integral del modelo de regresión y sus implicaciones.

Tabla 14*Construcción de las 63 combinaciones posibles entre las variables en estudio*

N°	Regresores	Coefs(<i>p</i>)	Regresores	Suma de cuadrados residuales SSRes(<i>p</i>)	Coefficiente de determinación múltiple R ²	R cuadrado ajustado	Cuadrado medio de residuales MSRes(<i>p</i>)	Cp	AIC	BIC
1	1	2	H/G	22680.23	0.40	0.13	810.01	30.13	202.84	205.64
2	1	2	H/DE	15926.37	0.64	0.39	568.80	13.42	192.24	195.04
3	1	2	H/PIB	17672.62	0.59	0.32	631.17	17.74	195.36	198.16
4	1	2	H/IDH	14876.90	0.67	0.43	531.32	10.82	190.19	192.99
5	1	2	H/PO	13328.74	0.71	0.49	476.03	6.99	186.89	189.70
6	1	2	H/GL	20925.58	0.47	0.20	747.34	25.79	200.43	203.23
7	2	3	H/G/DE	14686.15	0.67	0.41	543.93	12.35	191.80	196.01
8	2	3	H/G/PIB	12109.18	0.74	0.52	448.49	5.97	186.02	190.22
9	2	3	H/G/IDH	13011.51	0.72	0.48	481.91	8.20	188.17	192.38
10	2	3	H/G/PO	12496.01	0.73	0.50	462.82	6.93	186.96	191.16
11	2	3	H/G/GL	16688.23	0.62	0.33	618.08	17.30	195.64	199.84
12	2	3	H/DE/PIB	15892.03	0.64	0.37	588.59	15.33	194.17	198.37
13	2	3	H/DE/IDH	14176.82	0.69	0.43	0.63	11.09	190.74	194.95
14	2	3	H/DE/PO	12645.49	0.73	0.50	468.35	7.30	187.32	191.52
15	2	3	H/DE/GL	13097.63	0.72	0.74	485.10	8.42	188.37	192.57
16	2	3	H/PIB/IDH	14578.30	0.68	0.42	539.94	12.08	191.58	195.79
17	2	3	H/PIB/PO	13231.62	0.71	0.47	490.06	8.75	188.68	192.88
18	2	3	H/PIB/GL	14096.10	0.69	0.44	522.08	10.89	190.57	194.78
19	2	3	H/IDH/PO	13267.93	0.71	0.47	491.40	8.84	188.76	192.96
20	2	3	H/IDH/GL	13853.44	0.70	0.45	513.09	10.29	190.05	194.26
21	2	3	H/PO/GL	12472.96	0.73	0.50	461.96	6.87	186.90	191.11
22	3	4	H/G/DE/PIB	11133.77	0.77	0.54	428.22	5.56	185.50	191.10
23	3	4	H/G/DE/IDH	10419.18	0.78	0.57	400.74	3.79	183.51	189.11

24	3	4	H/G/DE/PO	9604.91	0.80	0.60	369.42	1.77	181.06	186.67
25	3	4	H/G/DE/GL	12796.23	0.72	0.47	492.16	9.67	189.67	195.28
26	3	4	H/G/PIB/IDH	11589.88	0.75	0.52	445.76	6.69	186.70	192.31
27	3	4	H/G/PIB/PO	11025.97	0.77	0.54	424.08	5.29	185.20	190.81
28	3	4	H/G/PIB/GL	11600.61	0.75	0.52	446.18	6.71	186.73	192.33
29	3	4	H/G/IDH/PO	11684.64	0.75	0.52	449.41	6.92	186.94	192.55
30	3	4	H/G/IDH/GL	13007.26	0.72	0.46	500.28	10.19	190.16	195.77
31	3	4	H/G/PO/GL	12215.65	0.74	0.49	469.83	8.23	188.28	193.88
32	3	4	H/DE/PIB/IDH	12283.81	0.74	0.49	472.45	8.40	188.45	194.05
33	3	4	H/DE/PIB/PO	10841.54	0.77	0.55	416.98	4.83	184.70	190.30
34	3	4	H/DE/PIB/GL	13018.72	0.72	0.46	500.72	10.22	190.19	195.79
35	3	4	H/DE/IDH/PO	12615.06	0.73	0.48	485.19	9.22	189.24	194.85
36	3	4	H/DE/IDH/GL	12664.80	0.73	0.48	487.11	9.35	189.36	194.97
37	3	4	H/DE/PO/GL	11542.43	0.76	0.52	443.94	6.57	186.58	192.18
38	3	4	H/PIB/IDH/PO	12816.48	0.72	0.47	492.94	9.72	189.72	195.32
39	3	4	H/PIB/IDH/GL	13816.14	0.70	0.43	531.39	12.20	191.97	197.58
40	3	4	H/PIB/PO/GL	12467.01	0.73	0.48	479.50	8.86	188.89	194.49
41	3	4	H/IDH/PO/GL	12454.17	0.73	0.48	479.01	8.83	188.86	194.46
42	4	5	H/G/DE/PIB/IDH	10263.04	0.79	0.56	410.52	5.40	185.05	192.06
43	4	5	H/G/DE/PIB/PO	9514.31	0.80	0.59	380.57	3.55	182.78	189.79
44	4	5	H/G/DE/PIB/GL	10749.97	0.78	0.54	430.00	6.61	186.44	193.45
45	4	5	H/G/DE/IDH/PO	9332.70	0.81	0.60	373.31	3.10	182.20	189.21
46	4	5	H/G/DE/IDH/GL	10391.60	0.78	0.55	415.66	5.72	185.43	192.43
47	4	5	H/G/DE/PO/GL	9590.83	0.80	0.59	383.63	3.74	183.02	190.03
48	4	5	H/G/PIB/IDH/PO	10947.98	0.77	0.53	437.92	7.10	186.99	194.00
49	4	5	H/G/PIB/IDH/GL	11487.90	0.76	0.51	459.52	8.43	188.44	195.44
50	4	5	H/G/PIB/PO/GL	10904.35	0.77	0.53	436.17	6.99	186.87	193.88
51	4	5	H/G/IDH/PO/GL	11684.64	0.75	0.50	467.37	8.92	188.94	195.95
52	4	5	H/DE/PIB/IDH/PO	10316.26	0.79	0.56	412.65	5.53	185.21	192.21
53	4	5	H/DE/PIB/IDH/GL	12156.37	0.74	0.48	486.25	10.09	190.13	197.14

54	4	5	H/DE/PIB/PO/GL	10605.08	0.78	0.54	424.20	6.25	186.04	193.04
55	4	5	H/DE/IDH/PO/GL	11374.23	0.76	0.51	454.97	8.15	188.14	195.14
56	4	5	H/PIB/IDH/PO/GL	12450.22	0.73	0.46	498.01	10.82	190.85	197.85
57	5	6	H/G/DE/PIB/IDH/PO	9331.63	0.81	0.58	388.82	5.10	184.20	192.61
58	5	6	H/G/DE/PIB/IDH/GL	10261.85	0.79	0.54	427.58	7.40	187.05	195.46
59	5	6	H/G/DE/PIB/PO/GL	9501.47	0.80	0.57	395.89	5.52	184.74	193.15
60	5	6	H/G/DE/IDH/PO/GL	9294.50	0.81	0.58	387.27	5.00	184.08	192.49
61	5	6	H/G/PIB/IDH/PO/GL	10895.85	0.77	0.51	453.99	8.97	188.85	197.26
62	5	6	H/DE/PIB/IDH/PO/GL	10316.10	0.79	0.54	429.84	7.53	187.21	195.62
63	6	7	H/G/DE/PIB/IDH/PO/GL	9292.64	0.81	0.57	404.03	7.00	186.07	195.88

Nota. La Tabla 14 muestra la construcción detallada de las 63 combinaciones posibles entre las variables regresoras que pueden interactuar con la variable respuesta. Cada celda de la tabla refleja la combinación específica de valores de las variables predictoras que se utiliza para evaluar su impacto conjunto en la variable de interés.

4.3.2 Selección de los modelos bajo los criterios Cp, AIC y BIC

Para construir el modelo que mejor se adapte al comportamiento de las variables bajo estudio, se optó por aplicar los tres criterios de selección mencionados anteriormente.

Para el Criterio de Información de Akaike (AIC) y el Criterio de Información Bayesiana (BIC), se ha seleccionado el valor más bajo obtenido de cada uno de los seis grupos diferentes de variables regresoras interactivas. Para el Estadístico Predictivo Conceptual de Mallows (Cp), se consideró el valor más bajo que se ajustara al número de variables regresoras de cada uno de los seis grupos diferentes de variables regresoras interactivas.

Retomando los resultados de la Tabla 13, en la que se calcularon los valores de AIC, BIC y Cp para las 63 combinaciones, se seleccionaron un total de 10 combinaciones utilizando estos tres criterios. Las combinaciones seleccionadas se presentan en la Tabla 15.

Tabla 15

Selección de modelos de acuerdo a los criterios Cp, AIC y BIC

N°	Regresores	Coefs(p)	Regresores	Suma de cuadrados residuales SSRes(p)	Coefficiente de determinación múltiple R ²	R cuadrado ajustado	Cuadrado medio de residuales MSRes(p)	Cp	AIC	SBC
5	1	2	H/PO	13328.738	0.711	0.488	476.026	6.989	186.894	189.697
8	2	3	H/G/PIB	12109.184	0.742	0.517	448.488	5.971	186.016	190.219
10	2	3	H/G/PO	12496.009	0.732	0.502	462.815	6.929	186.959	191.163
24	3	4	H/G/DE/PO	9604.910	0.802	0.602	369.420	1.773	181.065	186.670
29	3	4	H/G/IDH/PO	11684.639	0.753	0.516	449.409	6.920	186.945	192.550
45	4	5	H/G/DE/IDH/PO	9332.695	0.808	0.598	373.308	3.099	182.202	189.208
50	4	5	H/G/PIB/PO/GL	10904.350	0.772	0.530	436.174	6.989	186.872	193.878
58	5	6	H/G/DE/PIB/IDH/GL	10261.850	0.787	0.540	427.577	7.399	187.050	195.457
60	5	6	H/G/DE/IDH/PO/GL	9294.498	0.809	0.583	387.271	5.005	184.079	192.487
63	6	7	H/G/DE/PIB/IDH/PO/GL	9292.642	0.809	0.565	404.028	7.000	186.073	195.882

Nota. La Tabla 15 muestra las variables en estudio y el cálculo del Criterio de Información de Akaike (AIC), el Criterio de Información Bayesiano (BIC) y Estadístico Predictivo Conceptual de Mallows (Cp) que se han determinado con la aplicación adicional statistical Analysis Add-Ins For Excel (Completely Free) de cada uno de los 6 grupos de variables regresoras interactuando.

4.3.3 Propuesta de los Modelos Lineales

Tabla 16

Posibles Modelos seleccionados mediante el criterio AIC, BIC y Cp

N°	Regresores	Coefs(<i>p</i>)	Regresores	β_0	β_1	β_2	β_3	β_4	β_5	β_6
5	1	2	H/PO	-8.440	2.108					
8	2	3	H/G/PIB	598.509	-802.876	-0.057				
10	2	3	H/G/PO	41.507	-153.837	2.663				
24	3	4	H/G/DE/PO	94.434	-365.287	11.578	2.050			
29	3	4	H/G/IDH/PO	344.935	-254.498	-341.598	1.632			
45	4	5	H/G/DE/IDH/PO	271.222	-409.796	10.734	-203.373	1.480		
50	4	5	H/G/PIB/PO/GL	374.781	-534.586	-0.033	1.205	-0.454		
58	5	6	H/G/DE/PIB/IDH/GL	570.611	-569.370	9.595	-0.014	-399.012	0.058	
60	5	6	H/G/DE/IDH/PO/GL	311.793	-452.249	10.928	-252.527	1.488	0.309	
63	6	7	H/G/DE/PIB/IDH/PO/GL	307.514	-439.309	11.154	0.002	-267.980	1.513	0.336

Nota. La Tabla 16 muestra la propuesta de los coeficientes de los modelos seleccionados mediante el cálculo del Criterio de Información de Akaike (AIC), el Criterio de Información Bayesiano (BIC) y Estadístico Predictivo Conceptual de Mallows (Cp) de cada uno de los 6 grupos diferentes de variables regresoras interactuando.

Como ya se explicó, en el apartado anterior, no todas las variables cumplen con los cinco supuestos requeridos; acá se presentan los 10 mejores modelos de regresión que se utilizarían para las proyecciones, en caso que las variables en estudio cumplieran los supuestos. Sin embargo, estos modelos solo quedan indicados en este apartado.

En este sentido y debido al no cumplimiento de los supuestos, se recomienda la utilización de un modelo lineal generalizado. Tras analizar las características de estos modelos y las características de las variables en estudio, el modelo que mejor se adecua es el modelo de regresión de Poisson.

4.4 Modelo lineal generalizado

4.4.1 Modelo de Regresión de Poisson

La regresión log-lineal es uno de los casos especializados de modelos lineales generalizados para datos, este método se utiliza para modelar la relación entre una variable respuesta o dependiente y una o más variables explicativas o independientes, se encuentra disponible en Excel utilizando el complemento estadístico XLSTAT.

La variable respuesta o dependiente se escribe como el logaritmo de una función a fin de las variables explicativas o independientes. La regresión log-lineal en XLSTAT más común es la regresión de Poisson. Este enfoque se utiliza normalmente para modelar datos discretos.

Al contrario de la regresión lineal, no existe una solución analítica exacta. Por tanto, se debe utilizar un algoritmo iterativo. XLSTAT utiliza algoritmo de Newton-Raphson, se puede cambiar el número máximo de iteraciones y el umbral de convergencia si lo desea.

Los resultados de la regresión log-lineal en XLSTAT proporcionados son:

- a) Resumen de las estadísticas de las variables seleccionadas.
- b) Coeficientes de bondad de ajuste, en el que se puede observar una serie de estadísticos para el modelo independiente (correspondiente al caso donde la combinación lineal de variables explicativas se reduce a una constante) y para el modelo ajustado.
 - El número total de observaciones tomadas en cuenta (suma de los pesos de las observaciones)
 - Suma de ponderaciones o el número total de observaciones tomadas en cuenta (suma de las ponderaciones de las observaciones multiplicada por las ponderaciones en la regresión)
 - DF: Grados de libertad
 - -2 Log(Like.) , que es el logaritmo de la función de probabilidad asociada con el modelo
 - R^2 (McFadden), el coeficiente, al igual que el R^2 , se encuentra entre 0 y 1 que mide qué tan bien se ajusta el modelo. Este coeficiente es igual a 1 menos la relación entre la probabilidad del modelo ajustado y la probabilidad del modelo independiente;
 - R^2 (Cox y Snell), el coeficiente, al igual que el R^2 , se encuentra entre 0 y 1 que mide qué tan bien se ajusta el modelo. Este coeficiente es igual a 1 menos la relación entre

la probabilidad del modelo ajustado y la probabilidad del modelo independiente elevado a la potencia $2/S_w$, donde S_w es la suma de ponderaciones.

- R^2 (Nagelkerke): Coeficiente, al igual que el R^2 , entre 0 y 1 que mide qué tan bien se ajusta el modelo. Este coeficiente es igual a la relación del R^2 de Cox y Snell, dividido por 1 menos la probabilidad del modelo independiente elevado a la potencia $2/S_w$
- Desviación: criterio de desviación
- Chi-cuadrado de Pearson
- AIC, Criterio de Información de Akaike
- SBC, Criterio Bayesiano de Schwarz.

c) **Contraste de la hipótesis nula H_0 : $Y=constante$**

La hipótesis H_0 corresponde al modelo independiente que da el mismo resultado cualesquiera que sean los valores de las variables explicativas. Buscamos comprobar si el modelo ajustado es significativamente más potente que este modelo. Hay tres pruebas disponibles: la prueba de razón de verosimilitud (-2 Log (Like.)), la prueba de puntuación y la prueba de Wald. Las tres estadísticas siguen una distribución Chi^2 cuyos grados de libertad se muestran.

d) **Análisis tipo III**

Esta tabla sólo es útil si hay más de una variable explicativa. Aquí, el modelo ajustado se prueba con un modelo de prueba en el que se ha eliminado la variable de la fila de la tabla en cuestión. Si la resulta una probabilidad $\text{Pr} < 0.0001$, entonces la variable tiene información significativa, de lo contrario, se puede eliminar del modelo.

e) **Parámetros del modelo**

La estimación del parámetro, la desviación estándar correspondiente, el Chi^2 de Wald, el valor p correspondiente y el intervalo de confianza se muestran para la constante y cada variable del modelo.

f) **Ecuación del modelo**

Luego se muestra la ecuación del modelo para que sea más fácil de leer o reutilizar el modelo.

g) **Tabla de predicciones y residuos**

La tabla de predicciones y residuos muestra, para cada observación, su peso, el valor observado de la variable dependiente, la predicción del modelo, los mismos valores divididos por los pesos, los residuos estandarizados y un intervalo de confianza.

h) Prueba sobre dispersión

Para la regresión de Poisson se muestra una sobre dispersión.

4.4.2 Interpretación de los resultados

A continuación, se presentan los resultados obtenidos a partir de la aplicación de la regresión de Poisson. Estos resultados arrojaron un total de nueve modelos, los cuales se describen en las siguientes tablas y gráficos.

Modelo 1. Homicidios y Pobreza

Los resultados de las pruebas estadísticas del modelo 1 compuesto por la variable Homicidios y la variable Pobreza extrema, pueden observarse en la Tabla 17.

Tabla 17

Modelo 1. Parámetros del modelo compuesto por la variable Homicidios (H) y la variable Pobreza extrema (POBEXT)

Fuente	Valor	Error estándar	Chi-cuadrado de Wald	Pr > Chi ²	Wald Límite inf. (95%)	Wald Límite sup. (95%)
Intercepción	3.2658	0.0870	1410.4035	< 0.0001	3.0953	3.4362
POBEXT	0.0258	0.0020	173.6832	< 0.0001	0.0219	0.0296

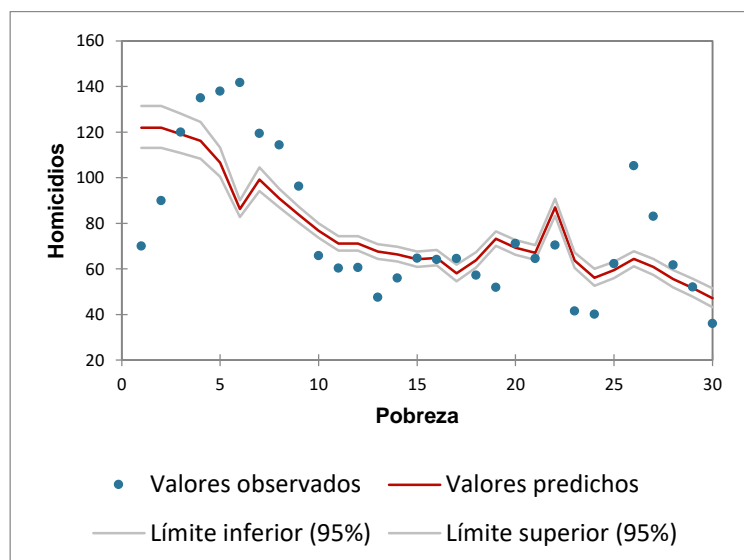
Nota. Puede observarse un p-valor de $P_r = < 0.0001$, lo que significa que la variable *POBEXT* es significativa.

El coeficiente de regresión de Poisson = 0.025, significa que por cada unidad que aumente la variable Pobreza, la tasa de incidencia de los Homicidios aumenta en un factor de aproximadamente 0.025%.

Además, el intervalo de confianza del 95% (IC95%) para este factor de aumento va desde 0.021 hasta 0.029. En otras palabras, existe una asociación estadísticamente significativa entre la variable pobreza y la variable homicidio, y este hallazgo sugiere que la pobreza está positivamente relacionada con la incidencia de la variable homicidios.

Figura 5

Comportamiento de valores predichos y observados para el modelo 1



Nota. La gráfica muestra la dispersión de los valores para el modelo 1, compuesto por la variable homicidios y la variable pobreza.

En este modelo compuesto por las variables antes mencionadas, de acuerdo a las pruebas estadísticas realizadas, puede concluirse que la variable independiente Pobreza es significativa en el crecimiento de los Homicidios.

Como resultado se obtuvo la ecuación del modelo 1: Homicidios-Pobreza, de la siguiente manera:

$$Pred(H) = e^{3.2658+0.0258 \cdot POBEXT}$$

Modelo 2. Homicidios, GINI y PIB

Los resultados de las pruebas estadísticas del modelo 2 compuesto por la variable Homicidios y las variables GINI y PIB, pueden observarse en la Tabla 18.

Tabla 18

Modelo 2. Parámetros del modelo compuesto por la variable Homicidios (H) y las variables GINI y PIB

Fuente	Valor	Error estándar	Chi-cuadrado de Wald	Pr > Chi²	Wald Límite inf. (95%)	Wald Límite sup. (95%)
Intercepción	10.2795	0.7118	208.5446	< 0.0001	8.8843	11.6746
GINI	-9.1529	1.2172	56.5449	< 0.0001	-11.5386	-6.7672
PIB	-0.0007	0.0001	123.8359	< 0.0001	-0.0008	-0.0005

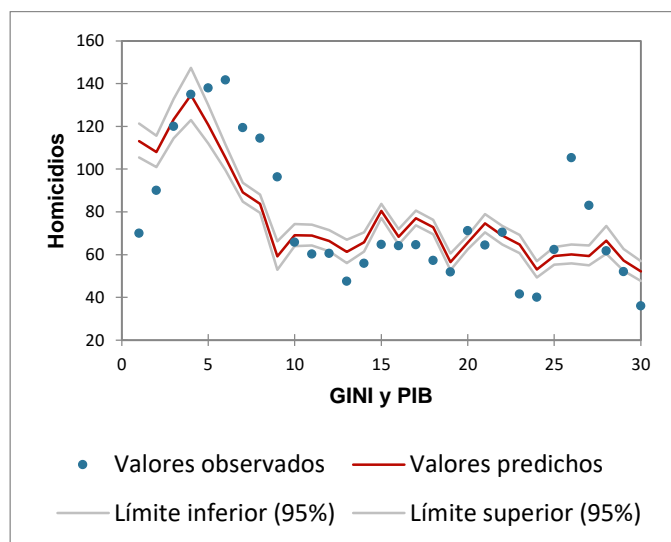
Nota. En la Tabla puede observarse que para ambas variables resultó un p-valor de $P_r < 0.0001$, lo que significa que las variables independientes *PIB* y *GINI* son significativas.

El coeficiente de regresión de Poisson indica que el valor para la variable independiente GINI es de -9.15 y para la variable independiente PIB es de -0.0007. Esto significa que, para la variable GINI, por cada unidad de aumento, la tasa de homicidios disminuirá en un 9.15%. En cuanto al PIB, por cada unidad de aumento en esta variable, hay una disminución del 0.0007% en la tasa de homicidios.

Además, el intervalo de confianza del 95% (IC95%) para este factor de aumento va desde -0.0008 hasta -0.0005.

Figura 6

Comportamiento de valores predichos y observados para el modelo 2



Nota. La gráfica muestra la dispersión de los valores para el modelo 2, compuesto por la variable homicidios y las variables GINI y PIB.

En este modelo compuesto por la variable dependiente Homicidios y las variables independientes GINI y PIB, de acuerdo a las pruebas estadísticas realizadas, puede concluirse que existe una asociación estadísticamente significativa y que de estas variables la que más aporta información al modelo es la variable GINI.

Como resultado se obtuvo la ecuación del modelo 2: Homicidios-GINI-PIB, de la siguiente manera:

$$Pred(H) = e^{10.2795 - 9.1529 \cdot GINI - 0.0007 \cdot PIB}$$

Modelo 3. Homicidios, GINI y Pobreza

Los resultados de las pruebas estadísticas del modelo 3 compuesto por la variable Homicidios y las variables GINI y Pobreza, pueden observarse en la Tabla 19.

Tabla 19

Modelo 3. Parámetros del modelo compuesto por la variable Homicidios (H) y las variables GINI y Pobreza

Fuente	Valor	Error estándar	Chi-cuadrado de Wald	Pr > Chi²	Wald Límite inf. (95%)	Wald Límite sup. (95%)
Intercepción	3.844	0.227	287.318	< 0.0001	3.399	4.288
GINI	-1.724	0.630	7.497	0.006	-2.958	-0.490
POBEXT	0.031	0.003	120.548	< 0.0001	0.026	0.037

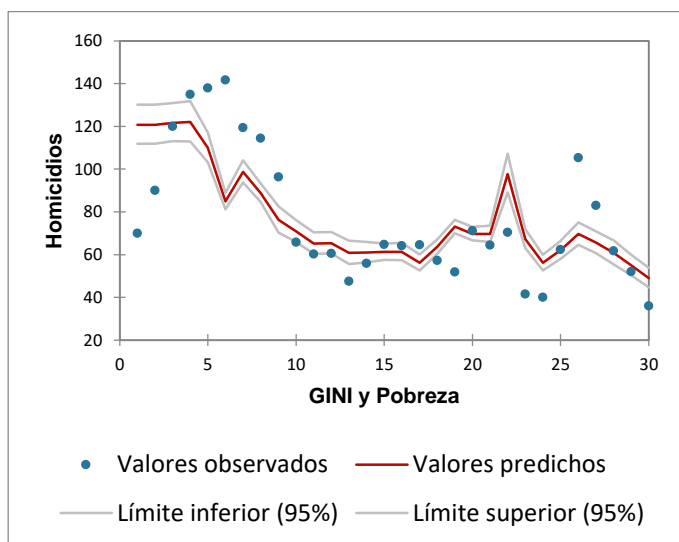
Nota. En la Tabla puede observarse que para la variable GINI resultó un p-valor de $Pr = 0.006$, por lo que para este modelo no es significativa. Sin embargo, la variable Pobreza (*POBEXT*) tiene un p-valor de $P_r < 0.001$, por lo que esta variable si es significativa para el modelo.

Este resultado indica que el valor del coeficiente de regresión de Poisson para la variable GINI es de -1.72 y para la Pobreza es de 0.031. Esto significa que, por cada unidad que aumente la variable GINI, la tasa de homicidios disminuirá en 1.72%.

En el caso de la variable Pobreza, por cada unidad de aumento en esta variable, hay un aumento del 0.031% en la tasa de homicidios.

Figura 7

Comportamiento de valores predichos y observados para el modelo 3



Nota. La gráfica muestra la dispersión de los valores para el modelo 3, compuesto por la variable homicidios y las variables GINI y Pobreza.

En este modelo compuesto por la variable dependiente Homicidios y las variables independientes GINI y Pobreza, de acuerdo a las pruebas estadísticas realizadas, puede concluirse que la variable (Desempleo) puede ser suprimida para este modelo 3, puesto que no tiene mayor incidencia en la variación de la tasa de homicidios.

Como resultado se obtuvo la ecuación del modelo 3: Homicidios-GINI-POBEXT, de la siguiente manera:

$$Pred(H) = e^{3.844 - 1.724 \cdot GINI - 0.031 \cdot POBEXT}$$

Modelo 4. Homicidios, GINI, DESEM (Desempleo) y POBEXT (Pobreza)

Los resultados de las pruebas estadísticas del modelo 4 compuesto por la variable Homicidios y las variables GINI, DESEM (Desempleo) y POBEXT (Pobreza), pueden observarse en la Tabla 20.

Tabla 20

Modelo 4. Parámetros del modelo compuesto por la variable Homicidios (H) y las variables GINI, DESEM (Desempleo) y POBEXT (Pobreza)

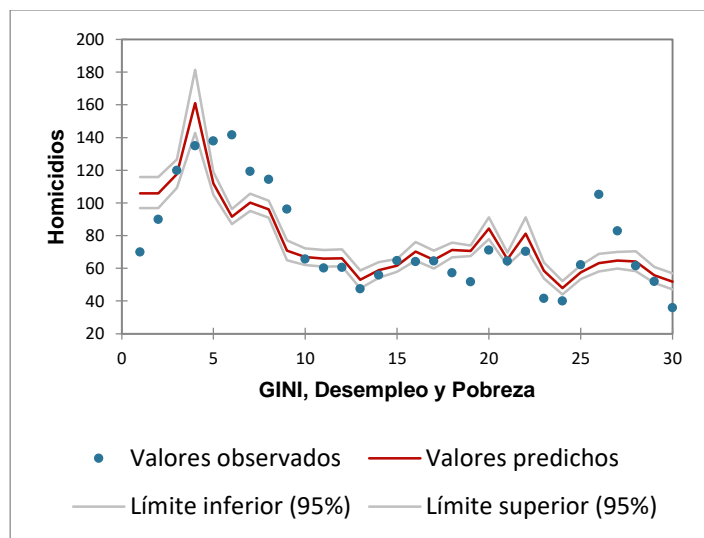
Fuente	Valor	Error estándar	Chi-cuadrado de Wald	Pr > Chi²	Wald Límite inf. (95%)	Wald Límite sup. (95%)
Intercepción	4.389	0.250	307.799	< 0.0001	3.899	4.880
GINI	-3.896	0.741	27.637	< 0.0001	-5.348	-2.443
DESEM	0.126	0.022	32.150	< 0.0001	0.083	0.170
POBEXT	0.024	0.003	55.439	< 0.0001	0.018	0.030

Nota. En la Tabla puede observarse que para todas las variables del modelo los p-valores obtenidos fueron igual a $P_r < 0.0001$, lo que significa que las variables independientes GINI, DESEM (Desempleo) y POBEXT (Pobreza), son significativas al modelo.

Este resultado indica que el valor del coeficiente de regresión de Poisson para la variable GINI es de -3.89, para el Desempleo 0.126 y para la Pobreza es de 0.024. Esto significa que, por cada unidad que aumente la variable GINI, la tasa de homicidios disminuirá en 3.89%; en el caso de la variable Desempleo, por cada unidad que aumente, la tasa de homicidios aumentará en 12.6%; mientras que, si la Pobreza aumenta en una unidad, la tasa de homicidios aumenta en un 0.024%.

Figura 8

Comportamiento de valores predichos y observados para el modelo 4



Nota. La gráfica muestra la dispersión de los valores para el modelo 4, compuesto por la variable homicidios y las variables GINI, DESEM (Desempleo) y POBEXT (Pobreza).

En este modelo compuesto por la variable dependiente Homicidios y las variables independientes GINI, DESEM (Desempleo) y POBEXT (Pobreza), de acuerdo a las pruebas estadísticas realizadas, puede concluirse que estas variables independientes son significativas en función del modelo.

Como resultado se obtuvo la ecuación del modelo 4: Homicidios-GINI-DESEM-POBEXT, de la siguiente manera:

$$Pred(H) = e^{4.389 - 3.896 \cdot GINI - 0.126 \cdot DESEM + 0.024 \cdot POBEXT}$$

Modelo 5. Homicidios, GINI, IDH y POBEXT (Pobreza)

Los resultados de las pruebas estadísticas del modelo 5 compuesto por la variable Homicidios y las variables GINI, IDH y POBEXT (Pobreza), pueden observarse en la Tabla 21.

Tabla 21

Modelo 5. Parámetros del modelo compuesto por la variable Homicidios (H) y las variables GINI, IDH y POBEXT (Pobreza)

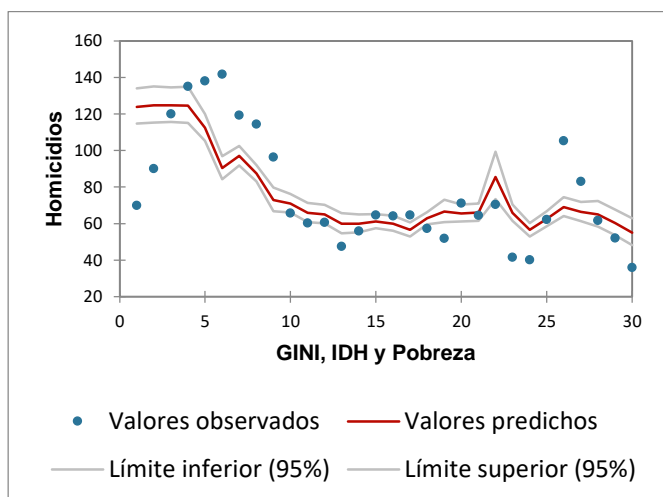
Fuente	Valor	Error estándar	Chi-cuadrado de Wald	Pr > Chi²	Wald Límite inf. (95%)	Wald Límite sup. (95%)
Intercepción	6.710	1.276	27.657	< 0.0001	4.209	9.211
GINI	-2.750	0.772	12.677	0.000	-4.263	-1.236
IDH	-3.160	1.390	5.167	0.023	-5.885	-0.435
POBEXT	0.022	0.005	16.867	< 0.0001	0.011	0.032

Nota. En la Tabla puede observarse que para la variable Pobreza resultó un p-valor de $P_r < 0.0001$, lo cual significa que contribuyen al ajuste del modelo. Sin embargo, las variables GINI e IDH resultó un p-valor de $P_r = 0.000$ y $P_r = 0.023$ respectivamente, por lo que estas dos variables no son significativas para este modelo.

Este resultado indica que el valor del coeficiente de regresión de Poisson para la variable GINI es de -2.750, para el IDH -3.160 y para la Pobreza es de 0.022. Esto significa que, por cada unidad que aumente la variable GINI, la tasa de homicidios disminuirá en 2.75%; en el caso de la variable IDH, por cada unidad que aumente, la tasa de homicidios disminuirá en 3.16%; mientras que, si la Pobreza aumenta en una unidad, la tasa de homicidios aumenta en un 0.022%.

Figura 9

Comportamiento de valores predichos y observados para el modelo 5



Nota. La gráfica muestra la dispersión de los valores para el modelo 5, compuesto por la variable homicidios y las variables GINI, IDH y POBEXT (Pobreza).

En este modelo compuesto por la variable dependiente Homicidios y las variables independientes GINI, IDH y Pobreza, de acuerdo a las pruebas estadísticas realizadas, puede concluirse que las variables IDH y POBEXT (Pobreza) puede ser suprimida para este modelo 5, puesto que no son significativas. Sin embargo, contribuyen a la disminución de la tasa de homicidios.

Como resultado se obtuvo la ecuación del modelo 5: Homicidios-GINI-IDH-POBEXT, de la siguiente manera:

$$Pred(H) = e^{6.710 - 2.750 \cdot GINI - 3.160 \cdot IDH - 0.022 \cdot POBEXT}$$

Modelo 6. Homicidios, GINI, DESEM (Desempleo), IDH y POBEXT

Los resultados de las pruebas estadísticas del modelo 6 compuesto por la variable Homicidios y las variables GINI, DESEM (Desempleo), IDH y POBEXT (Pobreza), pueden observarse en la Tabla 22.

Tabla 22

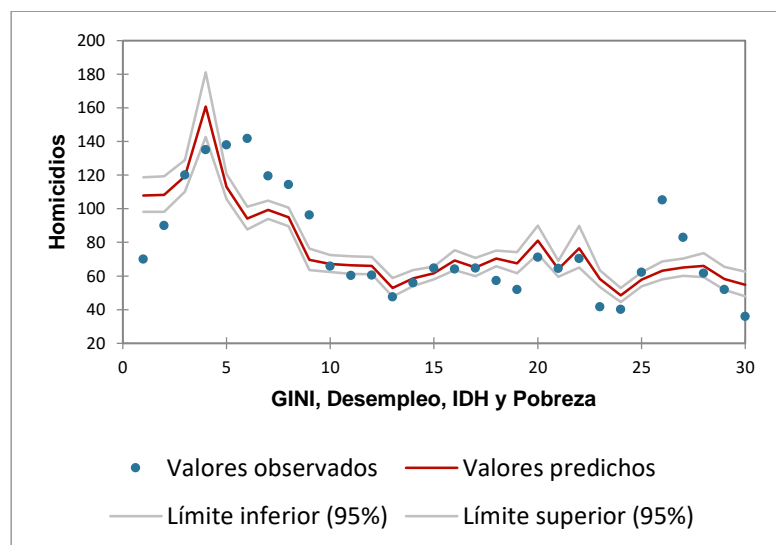
Modelo 6. Parámetros del modelo compuesto por la variable Homicidios (H) y las variables GINI, DESEM (Desempleo), IDH y POBEXT (Pobreza)

Fuente	Valor	Error estándar	Chi-cuadrado de Wald	Pr > Chi ²	Wald Límite inf. (95%)	Wald Límite sup. (95%)
Intercepción	5.769	1.299	19.718	< 0.0001	3.223	8.315
GINI	-4.287	0.823	27.147	< 0.0001	-5.899	-2.674
DESEM	0.121	0.023	28.104	< 0.0001	0.076	0.166
IDH	-1.555	1.440	1.168	0.280	-4.377	1.266
POBEXT	0.019	0.005	12.775	0.000	0.009	0.030

Nota. En la Tabla puede observarse que para las variables GINI y DESEM (Desempleo) resultó una probabilidad de $P_r < 0.0001$, las cuales contribuyen al ajuste del modelo. Sin embargo, la variable IDH y Pobreza no son significativas para este modelo.

Figura 10

Comportamiento de valores predichos y observados para el modelo 6



Nota. La gráfica muestra la dispersión de los valores para el modelo 6, compuesto por la variable Homicidio (H) y las variables GINI, DESEM (Desempleo), IDH y POBEXT (Pobreza).

En este modelo compuesto por la variable dependiente Homicidios y las variables independientes GINI, DESEM (Desempleo), IDH y POBEXT (Pobreza), de acuerdo a las pruebas estadísticas realizadas, puede concluirse que las variables IDH y POBEXT (Pobreza), pueden ser suprimida para este modelo 6, puesto que no es significativa.

Como resultado se obtuvo la ecuación del modelo 6: Homicidios-GINI-DESEM-IDH-POBEXT, de la siguiente manera:

$$\text{Pred}(H) = e^{5.769 - 4.287 \cdot \text{GINI} + 0.121 \cdot \text{DESEM} - 1.555 \cdot \text{IDH} - 0.019 \cdot \text{POBEXT}}$$

Modelo 7. Homicidios, GINI, PIB, POBEXT (Pobreza) y G&L (Gobierno de la Ley)

Los resultados de las pruebas estadísticas del modelo 7 compuesto por la variable Homicidios y las variables GINI, PIB, POBEXT (Pobreza) y G&L (Gobierno de la Ley), pueden observarse en la Tabla 23.

Tabla 23

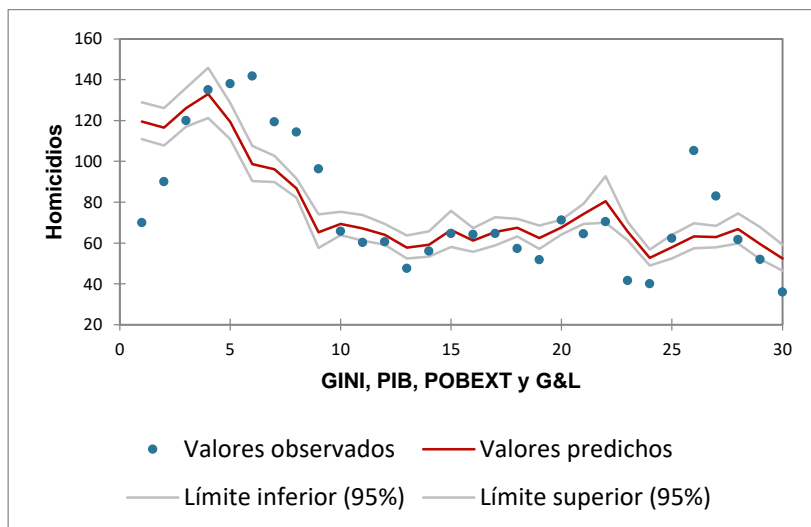
Modelo 7. Parámetros del modelo compuesto por la variable Homicidios (H) y las variables GINI, PIB, POBEXT (Pobreza) y G&L (Gobierno de la Ley)

Fuente	Valor	Error estándar	Chi-cuadrado de Wald	Pr > Chi²	Wald Límite inf. (95%)	Wald Límite sup. (95%)
Intercepción	7.619	1.047	52.911	< 0.0001	5.566	9.672
GINI	-5.958	1.545	14.879	0.000	-8.985	-2.931
PIB	0.000	0.000	12.626	0.000	-0.001	0.000
POBEXT	0.014	0.005	7.071	0.008	0.004	0.025
G&L	-0.006	0.005	1.257	0.262	-0.016	0.004

Nota. En la Tabla puede observarse que el P_r no es significativo para ninguna de las variables, por lo que *GINI, PIB, POBEXT* y *G&L* juntas no contribuyen significativamente al ajuste del modelo.

Figura 11

Comportamiento de valores predichos y observados para el modelo 7



Nota. La gráfica muestra la dispersión de los valores para el modelo 7, compuesto por la variable homicidios y las variables GINI, PIB, POBEXT (Pobreza) y G&L (Gobierno de la Ley).

En este modelo compuesto por la variable dependiente Homicidios y las variables independientes GINI, PIB, POBEXT (Pobreza) y G&L (Gobierno de la Ley), de acuerdo a las pruebas estadísticas realizadas, puede concluirse que estas variables independientes (GINI, POBEXT y G&L) puede ser suprimidas para este modelo 7, puesto que no son significativas.

Como resultado se obtuvo la ecuación del modelo 7: Homicidios-GINI-PIB-POBEXT-G&L, de la siguiente manera:

$$\text{Pred}(H) = e^{7.619 - 5.9583 \cdot \text{GINI} + 0.000 \cdot \text{PIB} + 0.014 \cdot \text{POBEXT} - 0.006 \cdot \text{G\&L}}$$

Modelo 8. Homicidios, GINI, DESEM (Desempleo), PIB, IDH y G&L (Gobierno de la Ley)

Los resultados de las pruebas estadísticas del modelo 8 compuesto por la variable Homicidios y las variables GINI, DESEM (Desempleo), PIB, IDH y G&L (Gobierno de la Ley), pueden observarse en la Tabla 24.

Tabla 24

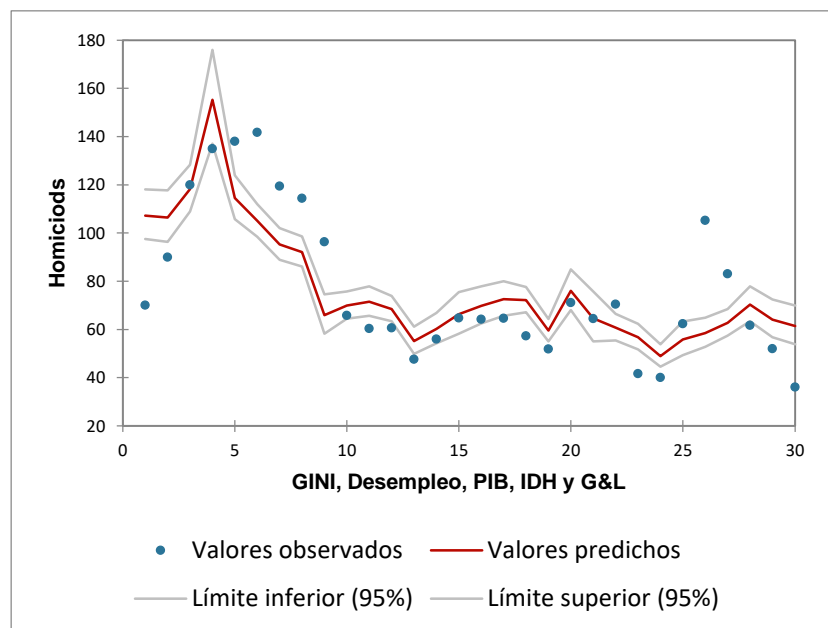
Modelo 8. Parámetros del modelo compuesto por la variable Homicidios (H) y las variables GINI, DESEM (Desempleo), PIB, IDH y G&L (Gobierno de la Ley)

Fuente	Valor	Error estándar	Chi-cuadrado de Wald	Pr > Chi ²	Wald Límite inf. (95%)	Wald Límite sup. (95%)
Intercepción	9.065	1.147	62.426	< 0.0001	6.816	11.314
GINI	-5.918	1.508	15.396	< 0.0001	-8.873	-2.962
DESEM	0.106	0.029	13.101	0.000	0.049	0.164
PIB	0.000	0.000	1.717	0.190	-0.001	0.000
IDH	-3.210	2.166	2.197	0.138	-7.456	1.035
G&L	-0.004	0.006	0.391	0.532	-0.017	0.009

Nota. En la Tabla puede observarse que la variable GINI tiene una probabilidad de $P_r < 0.0001$, lo cual significa que contribuye al ajuste del modelo. Sin embargo, las variables *DESEM* (Desempleo), *PIB*, *IDH* y *G&L* (Gobierno de la Ley) no son significativas para este modelo.

Figura 12

Comportamiento de valores predichos y observados para el modelo 8



Nota. La gráfica muestra la dispersión de los valores para el modelo 8, compuesto por la variable homicidios y las variables GINI, DESEM (Desempleo), PIB, IDH y G&L (Gobierno de la Ley).

En este modelo compuesto por la variable dependiente Homicidios y las variables independientes GINI, DESEM (Desempleo), PIB, IDH y G&L (Gobierno de la Ley), de acuerdo a las pruebas estadísticas realizadas, puede concluirse que las variables DESEM (Desempleo), PIB, IDH y G&L (Gobierno de la Ley) pueden ser suprimidas para este modelo 8, puesto que no son significativas.

Como resultado se obtuvo la ecuación del modelo 8: Homicidios-GINI-DESEM-PIB-IDH-G&L, de la siguiente manera:

$$Pred(H) = e^{9.065 - 5.918 \cdot GINI + 0.106 \cdot DESEM + 0.000 \cdot PIB - 3.210 \cdot IDH - 0.004 \cdot G\&L}$$

Modelo 9. Homicidios, GINI, DESEM (Desempleo), IDH, POBEXT (Pobreza) y G&L (Gobierno de la Ley)

Los resultados de las pruebas estadísticas del modelo 9 compuesto por la variable Homicidios y las variables GINI, DESEM (Desempleo), IDH, POBEXT (Pobreza) y G&L (Gobierno de la Ley), pueden observarse en la Tabla 25.

Tabla 25

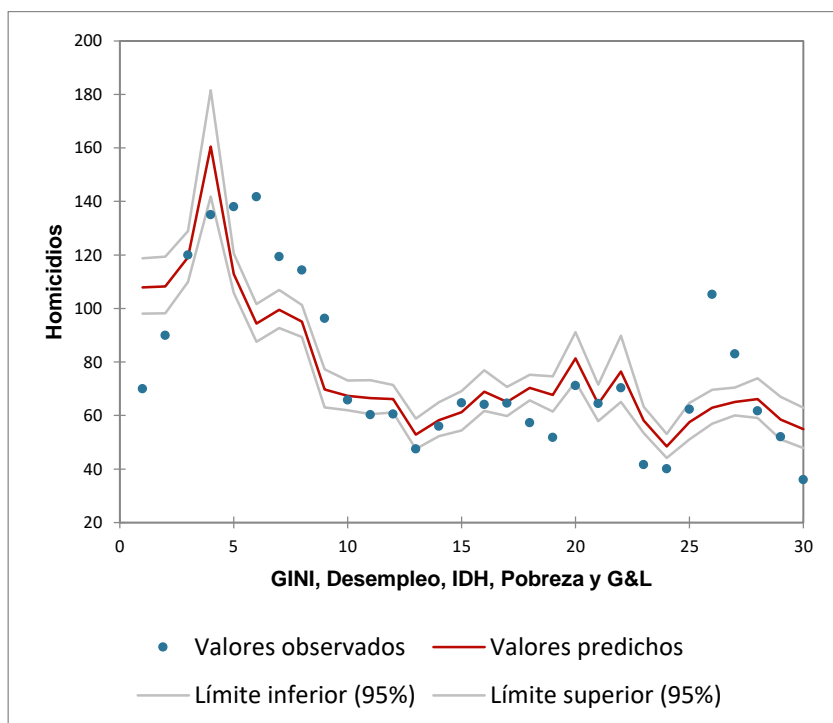
Modelo 9. Parámetros del modelo compuesto por la variable Homicidios (H) y las variables GINI, DESEM (Desempleo), IDH, POBEXT (Pobreza) y G&L (Gobierno de la Ley)

Fuente	Valor	Error estándar	Chi-cuadrado de Wald	Pr > Chi²	Wald Límite inf. (95%)	Wald Límite sup. (95%)
Intercepción	5.677	1.492	14.477	0.000	2.752	8.601
GINI	-4.194	1.108	14.334	0.000	-6.365	-2.023
DESEM	0.121	0.023	27.817	< 0.0001	0.076	0.166
IDH	-1.440	1.706	0.713	0.398	-4.784	1.903
POBEXT	0.019	0.005	12.750	0.000	0.009	0.030
G&L	-0.001	0.006	0.016	0.900	-0.013	0.011

Nota. En la Tabla puede observarse que únicamente la variable DESEM (Desempleo) resultó con una probabilidad de $P_r < 0.0001$, lo cual significa que contribuye al ajuste del modelo.

Figura 13

Comportamiento de valores predichos y observados para el modelo 9



Nota. La gráfica muestra la dispersión de los valores para el modelo 9, compuesto por la variable homicidios y las variables GINI, DESEM (Desempleo), IDH, POBEXT (Pobreza) y G&L (Gobierno de la Ley).

En este modelo compuesto por la variable dependiente Homicidios y las variables independientes GINI, DESEM (Desempleo), IDH, POBEXT (Pobreza) y G&L (Gobierno de la Ley), de acuerdo a las pruebas estadísticas realizadas, puede concluirse que las variables GINI, PIB, IDH, POBEXT (pobreza) y G&L (Gobierno de la Ley) pueden ser suprimidas para este modelo 9, puesto que no son significativas.

Como resultado se obtuvo la ecuación del modelo 9: Homicidios-GINI-DESEM-PIB-IDH-POBEXT-G&L (Gobierno de la Ley), de la siguiente manera:

$$Pred(H) = e^{5.677 - 4.194 \cdot GINI + 0.121 \cdot DESEM - 1.440 \cdot IDH + 0.019 \cdot POBEXT - 0.001 \cdot G\&L}$$

Modelo 10. Homicidios, GINI, DESEM (Desempleo), PIB, IDH, POBEXT (Pobreza) y G&L (Gobierno de la Ley)

Los resultados de las pruebas estadísticas del modelo 9 compuesto por la variable Homicidios y las variables GINI, DESEM (Desempleo), PIB, IDH, POBEXT (Pobreza) y G&L (Gobierno de la Ley), pueden observarse en la Tabla 26.

Tabla 26

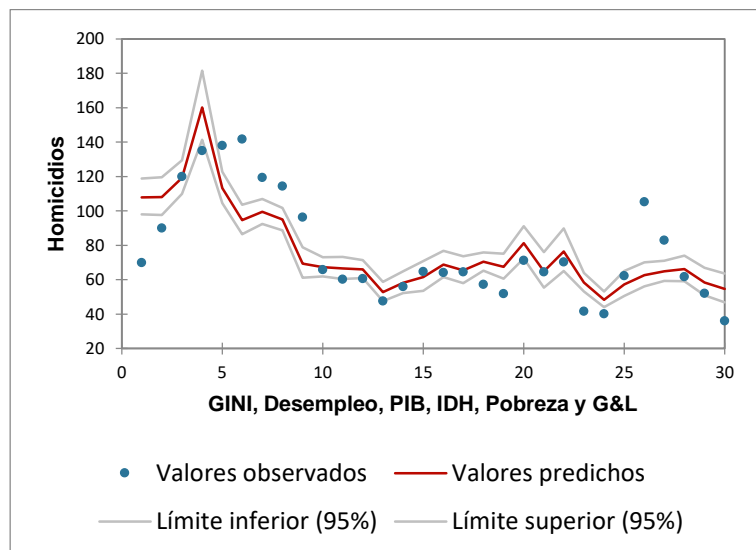
Modelo 10. Parámetros del modelo compuesto por la variable Homicidios (H) y las variables GINI, DESEM (Desempleo), PIB, IDH, POBEXT (Pobreza) y G&L (Gobierno de la Ley)

Fuente	Valor	Error estándar	Chi-cuadrado de Wald	Pr > Chi²	Wald Límite inf. (95%)	Wald Límite sup. (95%)
Intercepción	5.713	1.520	14.128	0.000	2.734	8.692
GINI	-4.337	1.581	7.521	0.006	-7.436	-1.237
DESEM	0.119	0.029	16.502	< 0.0001	0.061	0.176
PIB	0.000	0.000	0.016	0.899	0.000	0.000
IDH	-1.256	2.240	0.314	0.575	-5.646	3.134
POBEXT	0.019	0.006	10.956	0.001	0.008	0.030
G&L	-0.001	0.007	0.027	0.871	-0.014	0.012

Nota. En la Tabla puede observarse que únicamente la variable DESEM (Desempleo) resultó con una probabilidad de $P_r < 0.0001$, lo cual significa que contribuye al ajuste del modelo.

Figura 14

Comportamiento de valores predichos y observados para el modelo 10



Nota. La gráfica muestra la dispersión de los valores para el modelo 10, compuesto por la variable homicidios y las variables GINI, DESEM (Desempleo), PIB, IDH, POBEXT (Pobreza) y G&L (Gobierno de la Ley).

En este modelo compuesto por la variable dependiente Homicidios y las variables independientes GINI, DESEM (Desempleo), PIB, IDH, POBEXT (Pobreza) y G&L (Gobierno de la Ley), de acuerdo a las pruebas estadísticas realizadas, puede concluirse que las variables GINI, PIB, PIB, IDH, POBEXT (pobreza) y G&L (Gobierno de la Ley) pueden ser suprimidas para este modelo 10, puesto que no son significativas.

Como resultado se obtuvo la ecuación del modelo 10: Homicidios-GINI-DESEM-PIB-IDH-POBEXT-G&L (Gobierno de la Ley), de la siguiente manera:

$$\mathbf{Pred(H)} = e^{5.713 - 4.337 \cdot GINI + 0.119 \cdot DESEM + 0.000 \cdot PIB - 1.256 \cdot IDH + 0.019 \cdot POBEXT - 0.001 \cdot G\&L}$$

A continuación, se presentan un resumen de las ecuaciones de los modelos obtenido luego de la regresión de Poisson:

1. $Pred(H) = e^{3.2658+0.0258 \cdot POBEXT}$
2. $Pred(H) = e^{10.2795- 9.1529 \cdot GINI-0.0007 \cdot PIB}$
3. $Pred(H) = e^{3.844- 1.724 \cdot GINI+0.031 \cdot POBEXT}$
4. $Pred(H) = e^{4.389 - 3.896 \cdot GINI+0.126 \cdot DESEM+0.024 \cdot POBEXT}$
5. $Pred(H) = e^{6.710 -2.750 \cdot GINI-3.160 \cdot IDH+0.022 \cdot POBEXT}$
6. $Pred(H) = e^{5.769 -4.287 \cdot GINI+0.121 \cdot DESEM-1.555 \cdot IDH+0.019 \cdot POBEXT}$
7. $Pred(H) = e^{7.619 -5.9583 \cdot GINI+0.000 \cdot PIB+0.014 \cdot POBEXR-0.006 \cdot G\&L}$
8. $Pred(H) = e^{9.065 -5.918 \cdot GINI+0.106 \cdot DESEM+0.000 \cdot PIB -3.210 \cdot IDH-0.004 \cdot G\&L}$
9. $Pred(H) = e^{5.677 -4.194 \cdot GINI+0.121 \cdot DESEM-1.440 \cdot IDH+ 0.019 \cdot POBEXR-0.001 \cdot G\&L}$
10. $Pred(H) = e^{5.713 -4.337 \cdot GINI+0.119 \cdot DESEM+0.000 \cdot PIB - 1.256 \cdot IDH+0.019 \cdot POBEXR-0.001 \cdot G\&L}$

Con el objetivo de evaluar la funcionalidad de los modelos y determinar cuál predice mejor la variabilidad en la tasa de homicidios durante el período comprendido entre 1990 y 2019, se utilizaron los datos disponibles en la base de datos para cada una de las variables. Estos valores se sustituyeron en las ecuaciones correspondientes para realizar los cálculos año por año. Ver anexo 2.

Luego se aplicó el Error Cuadrático Medio (ECM), que mide la diferencia promedio entre los valores predichos y los reales. Ver anexo 3.

El ECM se calcula utilizando la fórmula:

$$ECM = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (30)$$

Donde y_i son los valores reales y \hat{y}_i son los valores predichos por el modelo.

Al realizar los cálculos, se determinó que el modelo con el menor Error Cuadrático Medio (ECM) fue el Modelo N° 2, cuya ecuación es:

$$Pred(H) = e^{10.2795- 9.1529 \cdot GINI-0.0007 \cdot PIB}$$

El segundo modelo con el menor ECM fue el Modelo N° 9, que se expresa de la siguiente manera:

$$Pred(H) = e^{5.677 - 4.194 \cdot GINI + 0.121 \cdot DESEM - 1.440 \cdot IDH + 0.019 \cdot POBEXR - 0.001 \cdot G\&L}$$

Finalmente, el tercer modelo con el menor ECM fue el Modelo N° 6, cuya ecuación es:

$$Pred(H) = e^{5.769 - 4.287 \cdot GINI + 0.121 \cdot DESEM - 1.555 \cdot IDH + 0.019 \cdot POBEXT}$$

Capítulo 5. Discusión

En esta investigación se analizaron un total de seis variables para determinar el grado de relación que existe con la tasa de homicidios en El Salvador en un periodo de tiempo que va de 1990 a 2019. Estas variables son: 1. Pobreza, 2. Desigualdad (GINI), 3. Desempleo, 4. Producto Interno Bruto per cápita (PIB), 5. Índice de Desarrollo humano (IDH) y 6. Institucionalidad (Gobierno de la Ley).

Cada una de estas variables, a excepción de la Institucionalidad, han sido utilizadas por diversos estudios para analizar y explicar la violencia en El Salvador. La naturaleza de estas variables es esencialmente económica y social.

Uno de los aportes de este trabajo de investigación es que se incorpora la variable **Institucionalidad (Gobierno de La Ley)** al análisis estadístico del comportamiento de la violencia en El Salvador y pretende comprobar si esta variable está asociada o no, con el aumento en la tasa de homicidios y si puede ser parte o no de un modelo estadístico que explique la incidencia de los homicidios en El Salvador.

Evidentemente, dentro del conjunto de variables consideradas en esta investigación, la variable "Gobierno de La Ley" muestra la segunda correlación más baja en relación con la tasa de homicidios. Sin embargo, es importante destacar que existe una correlación negativa significativa de -0.47, la cual no debe ser despreciada dentro del modelo propuesto. (Ver Tabla 11)

Una correlación de -0.47 indica una relación inversa moderada entre la variable "gobierno de la ley" y la tasa de homicidios. Esto significa que a medida que la calidad del gobierno de la ley mejora, es probable que la tasa de homicidios disminuya. Aunque la correlación no es tan fuerte como con otras variables, sigue siendo una relación importante a considerar en el análisis.

Es esencial recordar que la correlación no implica causalidad directa. No se puede afirmar que un mayor gobierno de la ley sea la causa única de una reducción en la tasa de homicidios. Sin embargo, la correlación negativa sugiere que hay una asociación entre estas dos variables, lo cual es relevante a considerar un modelo explicativo.

Es recomendable realizar un análisis más detallado y considerar otras variables para comprender mejor la relación entre el gobierno de la ley y la tasa de homicidios. Además, es importante tener en cuenta que este análisis se basa en los datos disponibles y puede estar sujeto a limitaciones y sesgos inherentes a la metodología utilizada y la calidad de los datos.

Los estudios realizados en El Salvador sobre el Gobierno de la Ley o la institucionalidad abordan el tema de manera general, sin un enfoque estadístico específico. Acevedo (2008) menciona las debilidades del marco institucional, mientras que Montalvo (2017) se centra en los factores contextuales e institucionales. Además, Carcach (2011) expresa la incapacidad de las instituciones encargadas del control del delito. Organismos internacionales como la OPS (2003) y la UNODC (2019) plantean que el fortalecimiento de la gobernanza y el estado de derecho tienen más probabilidades de ser efectivos para reducir el homicidio en América Latina.

Al analizar la correlación de las demás variables con la tasa de homicidios, se obtuvo que la que mayor correlación presenta es la **Pobreza** con 0.71; le sigue el **Índice de Desarrollo Humano (IDH)** con una correlación negativa de -0.66 ; la variable **Desempleo** presenta una correlación positiva de 0.64; el **Producto Interno Bruto (PIB)** presenta una correlación negativa de -0.58 ; mientras que la desigualdad medida a través del indicador **GINI**, presenta una correlación positiva de 0.39, siendo la que menor correlación presenta del conjunto de variables.

A continuación, se amplía el análisis de cada una de las variables mencionadas:

1. Pobreza Extrema. Esta variable muestra la correlación más alta en relación con la tasa de homicidios, con un valor de 0.71. Una correlación positiva tan fuerte indica que a medida que aumenta la pobreza extrema, es probable que también aumente la tasa de homicidios. Esto sugiere que la pobreza extrema puede ser un factor significativo en la incidencia de homicidios.

Este resultado coincide con el estudio realizado por Acevedo (2008) en el que se plantea que:

Los datos muestran una correlación apreciable entre las tasas de homicidios y las tasas de pobreza en el continente, pero en el caso de El Salvador –y de los países más violentos del continente– las tasas de homicidios se encuentran muy por encima de lo que sería de esperar dadas sus tasas de pobreza. (p. 74)

Esta correlación entre las tasas de homicidios y las tasas de pobreza en el continente latinoamericano ha sido documentada en diversos estudios. Estos estudios han demostrado que los países con mayores niveles de pobreza suelen tener tasas de homicidios más altas. Sin embargo, es importante señalar que esta relación no implica una causalidad directa y unidireccional. La pobreza extrema no es el único factor que explica la violencia, y existen otros elementos complejos y multifacéticos que influyen en la dinámica de la violencia en la región.

La pobreza extrema, entendida como la falta de acceso a recursos básicos, oportunidades educativas y empleo digno, crea un entorno propicio para la aparición y el aumento de la violencia. Las personas que viven en condiciones de pobreza a menudo se enfrentan a desigualdades estructurales, exclusión social y falta de perspectivas de desarrollo. Estos factores generan frustración, desesperanza y desconfianza en las instituciones, lo que puede impulsar la participación en actividades delictivas y violentas como una forma de supervivencia o búsqueda de poder.

Es fundamental comprender que la relación entre pobreza y violencia es bidireccional y se nutre de una serie de factores interrelacionados. La violencia puede contribuir a la perpetuación de la pobreza al generar inestabilidad económica, desplazamiento forzado, interrupción de servicios básicos y deterioro del tejido social. A su vez, la pobreza puede alimentar la violencia al crear condiciones propicias para la desigualdad, la exclusión y la falta de oportunidades.

2. Índice de Desarrollo Humano (IDH). El IDH presenta una correlación negativa de -0.66 con relación a la tasa de homicidios. Esta correlación indica que a medida que mejora el desarrollo humano, es probable que disminuya la tasa de homicidios. El IDH es una medida amplia que incluye indicadores como esperanza de vida, educación y nivel de vida. Esta correlación negativa sugiere que el desarrollo humano está asociado con una menor violencia homicida.

3. Desempleo. La variable desempleo muestra una correlación positiva de 0.64 con la tasa de homicidios. Esto implica que a medida que aumenta la tasa de desempleo, también puede aumentar la tasa de homicidios. El desempleo puede generar tensiones socioeconómicas y aumentar la desesperación en algunas comunidades, lo que podría contribuir a un aumento en la violencia.

En el estudio de Acevedo (2008), se destaca el desempleo juvenil como un factor social significativo que contribuye a la creación de un entorno propicio para la violencia. Específicamente, se enfoca en los jóvenes urbanos que han abandonado la escuela, quienes enfrentan mayores desafíos para acceder a empleo y oportunidades de desarrollo.

El desempleo juvenil, especialmente en contextos urbanos, puede tener múltiples implicaciones en términos de la incidencia de la violencia. La falta de empleo y perspectivas de

trabajo puede generar sentimientos de frustración, desesperanza y alienación entre los jóvenes, lo que aumenta su vulnerabilidad hacia involucrarse en actividades delictivas y violentas como una forma de subsistencia o búsqueda de reconocimiento y pertenencia en su entorno.

Además, el abandono escolar y la falta de educación formal pueden limitar las opciones de empleo y desarrollo personal de los jóvenes, aumentando aún más las tensiones y las posibilidades de involucrarse en comportamientos violentos. La falta de oportunidades económicas y sociales puede llevar a la marginalización de estos jóvenes en la sociedad, creando un caldo de cultivo para la violencia y la criminalidad.

Es importante tener en cuenta que el desempleo juvenil y el abandono escolar no actúan de manera aislada, sino que están interconectados con otros factores socioeconómicos y contextuales, como la falta de inversiones en educación, la escasez de empleos dignos, la discriminación estructural y las barreras para acceder a servicios básicos. Estos factores pueden intensificar los riesgos asociados con el desempleo juvenil y aumentar la probabilidad de que los jóvenes se vean involucrados en situaciones de violencia.

4. Producto Interno Bruto (PIB). Entre 1,990 y 2,020 el PIB muestra una correlación negativa de -0.58 con la tasa de homicidios. Esto indica que a medida que aumenta el PIB, es probable que disminuya la tasa de homicidios. El PIB es una medida del valor total de bienes y servicios producidos en un país, y su correlación negativa sugiere que un mayor nivel económico puede estar relacionado con una menor incidencia de homicidios.

Según el estudio de Carcach et al. (2011) en el que se compara el PIB per cápita y la tasa de homicidios, destacando periodos anteriores a 1992 y en los que se observa correlaciones entre estas variables:

Hasta 1969 y acorde a expectativas, la tasa de homicidio tendió a decrecer a medida que el PIB per Cápita crecía (Coeficiente de correlación igual a -0.68). Sin embargo, a partir de 1970, la tasa homicidio tendió ya sea a mantenerse estable, o a moverse en la misma dirección que el PIB per cápita (Coeficiente de correlación igual a 0.19). Un examen para sub-períodos dentro del período 1970-2008 indica que la correlación entre estas variables tiende a debilitarse de 0.93 durante el sub-período 1970-1979 a 0.24 durante 1980-1991, hasta alcanzar un valor pequeño y negativo de 0.07 para el sub-período 1992-2008. (p. 4)

Aunque este estudio concluye que los datos presentados no permiten determinar una relación causal entre el crecimiento económico y los homicidios. Esto sugiere que hay otras variables adicionales que influyen en la relación entre el crecimiento económico y la violencia, y que condicionan su evolución como la informalidad, que son aquellas actividades que escapan del control del gobierno como el contrabando, receptación de bienes ilegales, narcotráfico y comercio informal (Carcach et al., 2011).

5. Desigualdad (GINI). La desigualdad medida a través del indicador GINI presenta una correlación positiva de 0.39 con la tasa de homicidios. Esto sugiere que a medida que aumenta la desigualdad, también podría aumentar la tasa de homicidios. La desigualdad económica y social puede generar tensiones y resentimientos, lo que podría contribuir a una mayor incidencia de violencia homicida.

Regresión Lineal Simple

El análisis estadístico realizado se centró en examinar la relación entre la tasa de homicidios y diversas variables socioeconómicas. Se optó por utilizar el modelo de Regresión Lineal Simple, que es una técnica estadística que busca modelar la relación entre una variable dependiente y una variable independiente.

Los resultados obtenidos revelaron que la variable que presenta la correlación más fuerte con la tasa de homicidios es la pobreza. Esto significa que a medida que aumenta el nivel de pobreza, también tiende a aumentar la tasa de homicidios. La correlación Pearson de 0.71 indica una correlación positiva fuerte entre estas dos variables, lo que sugiere que existe una asociación significativa entre la pobreza y la incidencia de homicidios.

El coeficiente de determinación (R^2) de 0.50 proporciona información sobre la proporción de la variabilidad en la tasa de homicidios que puede ser explicada por la variable pobreza extrema en el modelo de regresión. En este caso, un R^2 de 0.50 significa que aproximadamente el 50% de la variabilidad en la tasa de homicidios puede ser explicada por la variable pobreza extrema, lo que indica que esta variable tiene un poder explicativo considerable sobre la incidencia de homicidios. Se considera que para muestras mayores a 20, “valores de r de menos de 0.5 indican por lo general poca correlación; valores de r cercanos o superiores a 0.85 indican una correlación lineal fuerte; mientras que superiores a 0.9 implican correlaciones muy fuertes” (de la Vara y Gutiérrez, 2012, p. 310).

Estos resultados sugieren que la pobreza extrema puede ser un factor importante a considerar al analizar y abordar el problema de la violencia homicida. Los hallazgos respaldan la idea de que las condiciones socioeconómicas precarias pueden contribuir significativamente al aumento de los homicidios.

Regresión Lineal Múltiple

La comprobación de los supuestos de normalidad, linealidad, independencia y homocedasticidad es crucial al aplicar el modelo de Regresión Lineal Múltiple. Estos supuestos son fundamentales para garantizar la validez de las inferencias realizadas a partir del modelo.

El supuesto de normalidad implica que las variables en estudio siguen una distribución normal. La linealidad supone una relación lineal entre las variables independientes y la variable dependiente. La independencia indica que no existe relación sistemática entre los residuos del modelo. La homocedasticidad implica que la varianza de los residuos es constante a lo largo de todos los niveles de las variables independientes.

Al realizar las pruebas estadísticas y comprobar los supuestos, se encontró que no todas las variables cumplían con ellos, lo que sugería que el modelo de Regresión Lineal Múltiple podría no ser adecuado en este caso.

Modelo Lineal Múltiple

Ante esta situación, se optó por realizar pruebas adicionales para determinar si el Modelo Lineal Múltiple era más apropiado. Este modelo permite incluir múltiples variables independientes en la ecuación de regresión, lo que podría ofrecer una mejor explicación de la variabilidad en la variable dependiente, incluso si no todas las variables cumplen con los supuestos mencionados.

La decisión de utilizar el Modelo Lineal Múltiple se basó en la consideración de que, a pesar de que algunas variables no cumplían con los supuestos, era importante incluirlas en el análisis debido a su relevancia teórica y su potencial relación con la variable dependiente.

Se obtuvieron 63 posibles modelos (aplicando $2^k - 1$), a los cuales se les aplicaron el Criterio de Información de Akaike (AIC), el Criterio de Información Bayesiano (BIC) y el Estadístico Predictivo Conceptual de Mallows (Cp). Luego de esta aplicación, se seleccionaron un total de 10 modelos basados en los valores más bajos de AIC y BIC, así como en el valor de Cp que se aproximara más al número de variables consideradas en el estudio.

Modelo lineal generalizado

Con los 10 modelos obtenidos, se sustituyeron los valores de cada una de las variables correspondientes al período de 1990 a 2019 en cada modelo, con el objetivo de observar cuál de ellos predice mejor la tasa de homicidios en El Salvador.

Este proceso consistió en reemplazar los valores históricos de las variables, los cuales se han registrado en la base de datos (ver anexo), en las ecuaciones de cada modelo para estimar la tasa de homicidios año por año durante ese período.

Posteriormente, se compararon las predicciones obtenidas de cada modelo con los datos reales de homicidios, evaluando el nivel de precisión y estimaciones más cercanas a los valores observados en la realidad.

Se seleccionó el modelo que mejor representaba el comportamiento histórico de la tasa de homicidios, lo que permitió identificar las variables más influyentes y su impacto sobre la violencia homicida en el país. El modelo está compuesto por las variables independientes: Desigualdad medido por GINI y Producto Interno Bruto (PIB) per cápita, y está representado por la siguiente ecuación de regresión:

$$Pred(H) = e^{10.2795 - 9.1529 \cdot GINI - 0.0007 \cdot PIB}$$

De acuerdo con el análisis de regresión GINI y el PIB per cápita son variables que presentan significancia dentro del modelo (ver Tabla 17). Por un lado, en cuanto al estadístico Chi Cuadrado, tanto la desigualdad (GINI), como el PIB tienen un valor $Pr > Chi^2$ inferior a 0.0001, lo que indica que ambos son altamente significativos. Esto significa que las variables GINI y PIB per cápita tienen un impacto estadísticamente significativo sobre la tasa de homicidios.

En cuanto al valor **Chi-cuadrado de Wald**, el cual mide la magnitud del efecto de cada variable sobre la variable dependiente (tasa de homicidios). En este caso, tanto el GINI (56.5449) como el PIB per cápita (123.8359) tienen valores elevados, lo que refuerza la idea de que estas variables tienen una influencia considerable en el modelo.

En lo que respecta al **error estándar** para GINI (1.2172) y el PIB per cápita (0.0001) es relativamente bajo, lo que sugiere que las estimaciones de los coeficientes son precisas. Un error estándar bajo implica mayor confianza en los valores estimados de los coeficientes de las variables, lo cual es favorable para la inclusión de estas variables en el modelo.

Los **límites de confianza de Wald** muestran un rango en el que se espera que caiga el verdadero valor del coeficiente con un 95% de confianza. En el caso del GINI, el intervalo va de -11.5386 a -6.7672, y para el PIB per cápita, de -0.0008 a -0.0005. Ambos intervalos no incluyen el valor 0, lo que refuerza que el efecto de estas variables es consistente y significativo.

En cuanto a la dirección del impacto de GINI (-9.1529) y PIB per cápita (-0.0007) son negativos, lo que indica que un aumento en el valor de estas variables está asociado con una disminución en la tasa de homicidios. Esto es consistente con la teoría en el caso del PIB, ya que un mayor crecimiento económico puede estar relacionado con una disminución de la violencia. En el caso del GINI, la interpretación sugiere que la desigualdad, medida a través de GINI, también tiene una relación inversa, esto difiere con la teoría ya que generalmente se espera que una mayor desigualdad (GINI) esté asociada con un aumento de la violencia.

Al realizar una comparación con el estudio estadístico sobre los homicidios que realizó Valiente (2019) en el cual aplicó el método de regresión cuantílica, obtuvo un modelo compuesto por 2 variables: cantidad de pandillas (Beta=16,62 y p=0,000) e índice de urbanidad (Beta=0,37 y p=0,000). Aunque los modelos difieren en las técnicas y las variables consideradas, ambos sugieren que ciertos factores sociales y económicos están asociados con la incidencia de homicidios en El Salvador, mostrando evidencia estadística.

Los escasos estudios estadísticos sobre los homicidios limitan la comparación de resultados obtenidos, lo que dificulta una comprensión más completa de este fenómeno. Sin embargo, otros estudios realizados en El Salvador, aunque no sean necesariamente de naturaleza estadística, respaldan la información utilizada sobre las variables, factores o determinantes de los homicidios en esta investigación. Esto permite una aproximación más profunda hacia su comprensión.

Desigualdad (GINI)

Si bien el fenómeno de la violencia tiene múltiples variables que la determinan, Acevedo (2008) plantea la siguiente agrupación para su explicación:

- i) Variables relacionadas con las condiciones generales de pobreza e inequidad; ii) Dificultades de acceso a educación y falta de oportunidades laborales, particularmente para la inserción de la población joven en el mercado de trabajo (tasas de desocupación y subempleo juvenil); iii) Factores asociados con la expansión del crimen organizado y

el fenómeno de las pandillas o maras; y iv) Debilidades del marco institucional para enfrentar el problema de la violencia con un enfoque preventivo efectivo. (p. 74)

La UNODC (2019) también se refiere a la desigualdad y plantea que entre mayor desigualdad mayores tasas de homicidio:

Los países con mayor desigualdad de ingresos tienen más probabilidades de tener tasas de homicidio más altas que los países con menor desigualdad. Los desarrollos económicos que exacerban la desigualdad de ingresos tanto dentro como entre países pueden fomentar la violencia criminal. (p. 15)

Como señala Carcach (2008) concluye que las desigualdades socioeconómicas son causas del aumento en las tasas de homicidios, “las tasas de homicidio son más altas en municipios con alta concentración de desventaja socioeconómica es válida” (p. 95).

En su estudio Valiente (2019), también plantea que “la desigualdad, medida por el coeficiente de Gini y la pobreza son predictores significativos de la tasa de homicidios a nivel mundial” (p. 152).

PIB per cápita

De acuerdo con Carcach et al. (2011), el crecimiento económico no es una causa de los homicidios. Sin embargo, correlaciona el PIB per cápita y la tasa de homicidios (p. 5). Aunque este estudio no profundiza en otros elementos estadísticos, sugiere no solo al PIB per cápita como una variable a considerar en el análisis de los homicidios, sino también otras como la informalidad.

Por otra parte, el Informe Mundial Sobre La Violencia y La Salud de la OPS expresa que “la tasa de crecimiento del producto interno bruto (PIB) también se asoció significativamente con una reducción de la tasa de homicidios” (p. 41).

Por la característica anterior en la que las tasas de homicidios tienen un comportamiento muy por encima de lo que aumentan o disminuyen las variables, la variable que mejor puede explicar este comportamiento es “la inequidad social, combinada perversamente con otras condiciones económicas, políticas y sociales, lo que explica las mayores expresiones de violencia en América Latina” (Acevedo, 2008, p. 74).

Por otra parte, Carcach (2011) plantea que “la problemática de la violencia se ha asociado a factores de carácter estructural tales como pobreza, desigualdad, polarización y exclusión”. En 2008 también señaló que “bajos niveles de desarrollo económico implican altos grados de

pobreza y niveles igualmente altos de actividad criminal” (Carcach, 2008, p. 10), relacionando de esta manera la pobreza con la criminalidad.

La variable pobreza también es considerada como una causa de la tasa de homicidios, tal como lo señala UNODC (2019) “existen otros factores, como la pobreza y la marginación, que influyen en la tasa de homicidios” (p. 32).

Conclusiones

Al analizar la correlación de las variables con la tasa de homicidios en el periodo comprendido entre 1990 y 2019, que abarca tres décadas de estudio, se observa que todas las variables consideradas en esta investigación muestran algún grado de correlación. Sin embargo, esta correlación no es uniforme, ya que se pueden identificar diferentes niveles de asociación entre estas variables. Algunas de ellas exhiben correlaciones notablemente más fuertes, sugiriendo una influencia más directa o determinante en la dinámica de la tasa de homicidios, mientras que otras presentan correlaciones más débiles, indicando una relación menos directa o más compleja con respecto a la variación en la tasa de homicidios a lo largo del tiempo. Este espectro de asociaciones nos proporciona una visión matizada y multifacética de los factores que influyen en la dinámica de la violencia homicida, destacando la importancia de un análisis detallado y contextualizado de cada variable para comprender adecuadamente su contribución al fenómeno de la criminalidad violenta.

Dadas las características estadísticas de las variables en estudio, se determinó que el tipo de regresión más adecuado sería la de Poisson. La elección de este modelo se fundamenta en su capacidad para manejar datos discretos, como es el caso de la tasa de homicidios, y para modelar adecuadamente la relación entre variables explicativas y la variable de interés.

El modelo de regresión de Poisson aplicado en este estudio se compone de dos variables que, en conjunto, mostraron una significancia estadística más alta. Estas variables incluyen la Desigualdad (medida por el índice GINI) y el PIB per cápita. Cada una de estas variables ofrece una perspectiva única sobre los factores socioeconómicos que pueden estar asociados con la incidencia de homicidios en la población.

La desigualdad económica, medida a través del índice GINI, proporciona información sobre la disparidad en la distribución de ingresos dentro de la sociedad, lo que puede tener implicaciones importantes para la cohesión social y la percepción de injusticia. Por su parte, el PIB per cápita que representa el crecimiento económico, refleja el nivel de riqueza o desarrollo económico medio de las personas. Juntas, estas variables ofrecen una visión integral de los factores socioeconómicos subyacentes que pueden contribuir a la incidencia de homicidios en El Salvador.

Aunque la variable "Gobierno de la Ley" no está incluida en el modelo propuesto, es importante destacar su relevancia, ya que muestra una correlación negativa con la tasa de

homicidios. Esta observación subraya la importancia del estado de derecho y la aplicación efectiva de las leyes en la reducción de la violencia homicida. La correlación negativa sugiere que a medida que la presencia o la fortaleza de un gobierno que aplica el estado de derecho aumenta, la tasa de homicidios tiende a disminuir. Este hallazgo es coherente con la teoría y la evidencia empírica que respaldan la idea de que un sistema legal sólido y una administración imparcial de la justicia son elementos cruciales para mantener la seguridad y el orden social.

Aunque esta variable no se ha incluido en el modelo estadístico propuesto, debido a que no cumple con los valores estimados de las diferentes pruebas estadísticas que se han aplicado, su correlación negativa ofrece una perspectiva valiosa sobre los factores que podrían influir en la incidencia de homicidios. La evidencia de esta correlación sugiere que el fortalecimiento de las instituciones gubernamentales encargadas de hacer cumplir la ley y garantizar la justicia puede desempeñar un papel significativo en la reducción de la violencia homicida. Esta observación subraya la importancia de considerar no solo los factores socioeconómicos, como la pobreza y el desempleo, sino también las dimensiones institucionales y gubernamentales en el análisis de la criminalidad y la violencia en una sociedad determinada.

Finalmente, de acuerdo con los resultados de la regresión de Poisson la variable GINI es significativo y negativo en varios modelos (como en los Modelos 2, 3, 4, 5, 6). La Pobreza extrema (POBEXT) y desempleo (DESEM) son positivas y significativas en la mayoría de los modelos donde aparecen. Por su parte, el impacto del PIB per cápita en los modelos tiene un impacto menor o no significativo en algunos casos. En este contexto, GINI y la pobreza parecen tener mayor relevancia en la predicción de los homicidios que el PIB per cápita.

Recomendaciones

A partir del proceso investigativo se han identificado principalmente tres recomendaciones generales: 1. Registro y automatización de los datos; 2. Selección y análisis estadístico de nuevas variables; y 3. Robustecer el modelo propuesto con nuevas variables.

El registro y automatización de datos, presentan una dificultad en este tipo de investigación, que requiere la creación de modelos para explicar el comportamiento de fenómenos sociales, económicos, de salud, medioambientales, entre otros. El desinterés de las autoridades en el registro de la información ha sido un obstáculo que impide disponer de los datos necesarios para realizar análisis estadísticos adecuados, a pesar de los avances tecnológicos y la facilidad que proporcionan los sistemas informáticos actuales para realizar estos análisis, que podrían contribuir al desarrollo de la sociedad.

Es en este sentido que no ha habido una cultura de salvaguardar los datos, ni de registrarlos a nivel informático, lo que ha resultado en la casi inexistencia de registros históricos. Por lo tanto, es recomendable iniciar un proceso de automatización de datos para la creación de una base de datos histórica de nuestro país. Esta base de datos debería permitir el acceso a información de diferentes variables económicas, sociales, políticas, de salud, medioambientales, entre otras. Estos registros históricos permitirían realizar mejores análisis y proyecciones, y proporcionar respuestas más acertadas a las diversas problemáticas, basándose en las herramientas informáticas y estadísticas actuales.

En cuanto a la selección y análisis de nuevas variables, al considerar más variables relevantes, se pueden identificar posibles relaciones y patrones adicionales que contribuyan a explicar las variaciones en la tasa de homicidios. Estas variables pueden incluir aspectos demográficos, políticos, culturales, entre otros. Por ejemplo, se podrían considerar variables como el acceso a la educación, la densidad de población, la presencia de grupos criminales y de pandillas (como en el estudio de Valiente, 2019), y políticas de seguridad. Otros factores, como menciona Carcach et al. (2011), son el narcotráfico, la corrupción y la economía informal.

La inclusión de nuevas variables puede ayudar a identificar otros factores que influyen en la violencia y los homicidios. Además, se pueden crear variables derivadas o combinaciones de variables existentes para obtener métricas más específicas o representativas. Por ejemplo, se podrían calcular índices propios a partir de las variables existentes y que ya han sido analizadas en esta investigación.

Dentro de esta recomendación, es importante tener cuidado al agregar nuevas variables, ya que no todas pueden ser relevantes o tener un impacto significativo en el modelo. Es necesario realizar un análisis cuidadoso, considerando la calidad de los datos disponibles, la teoría que las respalda y la coherencia conceptual al seleccionar las variables a incorporar.

Robustecer el modelo siempre será una alternativa, incorporar más variables puede ser beneficioso para corroborar el comportamiento de la tasa de homicidios y ampliar el análisis estadístico. Al agregar nuevas variables, se obtiene una visión más completa y precisa de los factores que influyen en la violencia medida a través de los homicidios.

Finalmente, considerando todo lo anterior, es necesario seguir revisando y afinando los modelos propuestos, de manera que se puedan realizar más pruebas estadísticas que permitan mejorar la precisión en la predicción de la tasa de homicidios

Referencias

- Abalos, E. y Rebollo, P. (2022). *Metodología de la investigación-recopilación*. Autores de Argentina.
- Acevedo, C. (2008). Los costos económicos de la violencia en El Salvador. *América Latina Hoy*, 50. <https://doi.org/10.14201/alh.1341>
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19, pp. 716–723.
- Aldrich, J. (1997). R.A. Fisher and the making of maximum likelihood 1912–1922. *Statistical Science*, 12(3), pp. 162–176.
- Anceschi, A. (2009) *La violencia familiar: aspetti penali, civil e criminologici*. Torino: G. Giappichelli.
- Anderson, D. R., Sweeney, D. J., Williams, T. A., Roa, M. D. C. H., & Álvarez, T. L. (2001). *Essentials Modern Business Statistics*. Cengage Learning Editores, S.A de C.V.
- Anderson, D., & Sweeney, D. (1999). *Estadística para administración y economía*. International Thomson Editores.
- Banco Interamericano de Desarrollo. (1998). *Facing Up to Inequality in Latin America. Economic and Social Progress in Latin America Report*.
- Banco Mundial [BM]. (2021). *Worldwide Governance indicators*. <https://databank.worldbank.org/source/worldwide-governance-indicators>
- Beltrán, M. (2000). *Perspectivas sociales y conocimiento*. Anthropos.
- Brender, N. (2012). *Investigando el dilema urbano: urbanización, pobreza y violencia. Resumen*. Centro Internacional de Investigaciones para el Desarrollo.
- Briceño-León, R. (2012). La Comprensión de los Homicidios en América Latina: ¿Pobreza o Institucionalidad? *Ciência & Saúde Coletiva*, 17, 3159-3170. <https://doi.org/10.1590/S1413-81232012001200002>
- Briceño-León, R., Ávila, O. & Camardiel, A. (2012). *Violencia e institucionalidad. Informe del observatorio venezolano de violencia 2011*. Editorial Alfa.
- Buvinic, M., Morrison, A. & Orlando, M. (2005). Violencia, crimen y desarrollo social en América Latina y el Caribe. *Papeles de población*, 43, 167–214.

- Buvinic, M., Morrison, A. & Shifter, M. (1999). *La violencia en América Latina y el Caribe: un marco de referencia para la acción*. Banco Interamericano de Desarrollo.
- Calabrese, E. (1997). La Violencia en el hogar. *Leviatán, Revista de hechos e ideas, II Época, n° 69*.
- Cameron, A. (1998). *Regression analysis of count data*.
https://assets.cambridge.org/97805216/32010/frontmatter/9780521632010_frontmatter.pdf
- Canavos, G. (1987). *Probabilidad y estadística*. McGraw Hill.
- Carcach, C. (2008). *El Salvador. Mapa de violencia y su referencia histórica*. Centro de Monitoreo y Evaluación de la Violencia desde la Perspectiva Ciudadana.
- Carcach, C., Rivera, J. E., y Morera, L. (2011). *Violencia y Crecimiento Económico*.
- Comisión Económica para América Latina y el Caribe [CEPAL]. (2019). *Panorama Social de América Latina*, (LC/PUB.2019/22-P/Re v. 1).
- Comisión Económica para América Latina y el Caribe [CEPAL] y la Organización Internacional del Trabajo [OIT]. (2019). Crisis en los mercados laborales y respuestas contracíclicas. *Coyuntura laboral en América Latina y el Caribe*, 2, 1-20.
- Comisión Interamericana de Derechos Humanos [CIDH]. (2021). *Informe de país. Situación de los derechos humanos en El Salvador*. Organización de Estados Americanos (OEA)
- De la Vara, R. & Gutiérrez, H. (2012). *Análisis y diseños de experimentos*. (3ª Edición). McGraw-Hill.
- Dirección General de Estadística y Censos [DIGESTYC]. (2019). *Encuesta de Hogares de Propósitos Múltiples 2018*. El Salvador: Ministerio de Economía.
- Duque, L. F. y Klevens, J. (2000). Creencias, Actitudes y Práctica asociadas con la violencia en Bogotá. *Coyuntura Social*, 22. Bogotá: Fedesarrollo e Instituto SER de Investigación.
- Echeverría, J. (1999). *Introducción a la metodología de la ciencia. La filosofía de la ciencia en el SXX*. Madrid: CATEDRA. Obtenido de Dialnet-
MetodosYTecnicasDeInvestigacionUtilizadosEnLosEstu-7706797.pdf
- Elorza, H. & Medina, J. (1999). *Estadística para las ciencias sociales y del comportamiento*. (3ra Edición). Oxford University.
- Everitt, B. (2002). *The Cambridge Dictionary of Statistics*. Cambridge University Press.

- Fajnzlber, P., Lederman, D. & Loayza, N. (2002). Inequality and violent crime. *Journal of Law and Economics*, 45, 1–40.
- Fernández, C., Hernández, R., y Baptista P. (2014). *Metodología de la Investigación (Sexta edición)*. McGraw-Hill.
- Fernández, C., Hernández, R., y Baptista P. (2006). *Metodología de la Investigación (Cuarta edición)*. McGraw-Hill.
- Fundación de Estudios para la Aplicación del Derecho [FESPAD]. (2021). La desaparición de personas y el contexto de violencia actual en El Salvador: una aproximación inicial.
- Ganchozo, B., Hernández, A., Moreno, L., Placencia B., Quimis, A. y Ramos, M. (2018). *Metodología de la investigación científica*. Editorial Área de Innovación y Desarrollo, S.L.
- Galtung, J. (1998). *Tras la violencia, 3R: reconstrucción, reconciliación, resolución*. Bakeas.
- Gill, J. (2001). *Generalized linear models: a unified approach*. Sage Publications.
- González, Á. R. A., Escobar-Córdoba, F., & Castañeda, G. C. (2007). Factores de riesgo para violencia y homicidio juvenil. *Revista Colombiana de Psiquiatría*, 1, 20.
- González-Pérez, G., Vega-López, M., Cabrera-Pivaral, C., Vega-López, A. & Muñoz de la Torre, A. (2012). Mortalidad por homicidios en México: Tendencias, variaciones socio-geográficas y factores asociados. *Ciência & Saúde Coletiva*, 17, 3195-3208. <https://doi.org/10.1590/S1413-81232012001200005>
- Gorgas, J. G., Cardiel, N. L. & Zamorano, J. C. (2011). *Estadística Básica para Estudiantes de Ciencias*. Universidad Complutense de Madrid.
- Hernández-Sampieri, R. (1998). *Metodología de la investigación*. McGraw-Hill.
- INTERPEACE y Programa Políticas Públicas para Prevenir la Violencia Juvenil (POLJUVE). (2009). *Violencia Juvenil, Maras y Pandillas. Informe para la Discusión*.
- Johnson, R. (1999). *Estadística elemental (Lo esencial) (2ª Edición)*. Ediciones Paraninfo.
- Kline, R.B. (1998). *Principles and practice of structural equation modeling*. The Guilford Press.
- La Parra, D. & Tortosa, J. M. (2003). Violencia estructural: una ilustración del concepto. *Documentación social*, 131(3), 57-72.

- Levin, R., Rubin, D., Rastogi, S. & Hussain, M. (2014). *Statistics for Management*. Pearson.
<https://www.amazon.in/Statistics-Management-7e-IEVIN-Rastogi/dp/8131774503>
- Long, J. (1997). *Regression models for categorical and limited dependent variables*. Sage Publications.
- Makridakis, S., Wheelwright, S. C. & Hyndman, R. J. (1978). *Forecasting methods and applications*. 2a edición. Wiley.
- Martínez, R., Tuya, L. & Pérez, A. (2009). El coeficiente de correlación de los rangos de Spearman. *Revista Habanera de Ciencias Médicas*, 99-100.
- Martínez, R. (2019). *Institucionalidad social en América Latina y el Caribe*. Comisión Económica para América Latina y El Caribe.
- Mendenhall, III (William), Beaver, R. J., & Beaver, B. M. (2017). *Probabilidad y estadística para las ciencias sociales del comportamiento y la salud*. Cengage Learning Editores, S.A de C.V.
- Montalvo, N. (2017). *Módulo I. Marco teórico y conceptual de la violencia*. Fundación Nacional para el Desarrollo. <https://repo.funde.org/1243/>
- Montgomery, D., Peck, E. & Vining, G. (2006). *Introducción al análisis de regresión lineal*. Compañía editorial comercial.
- Muñoz, C. (2015). *Metodología de la Investigación*. Oxford University Press.
- Novalés, A. (2010). *Análisis de regresión*. Universidad Complutense de Madrid.
[https://www.ucm.es/data/cont/docs/518-2013-11-13-Analisis de Regresion.pdf](https://www.ucm.es/data/cont/docs/518-2013-11-13-Analisis%20de%20Regresion.pdf)
- Oficina de las Naciones Unidas contra la Droga y el Delito [UNODC]. (2019). *Estudio Mundial sobre el Homicidio*. Resumen Ejecutivo.
- Olasolo, A., Pérez, M. & Ruiz, V. (2008). Análisis de la eficacia de las medias móviles en el mercado intradiario de renta variable español. *Universidad, Sociedad y Mercados Globales*. 56-68.
<https://dialnet.unirioja.es/servlet/articulo?codigo=2747743>
- Oliver, A., Rodrigo, M. & Tomás, J. (2005). Modelos lineales y no lineales en la explicación de la siniestralidad laboral. *Psicothema*, 17(1), 154-163.

- Organización Mundial de la Salud [OMS]. (2014). *Informe sobre la situación mundial de la prevención de la violencia 2014*. Organización Mundial de la Salud.
- Organización Mundial de la Salud. [OMS]. (2002). *Informe Mundial Sobre la Violencia y la Salud*. Sinopsis.
- Organización Panamericana de la Salud, Oficina Regional para las Américas de la Organización Mundial de la Salud. (2003). *Informe Mundial Sobre la Violencia y la Salud*.
- Programa de las Naciones Unidas para el Desarrollo [PNUD]. (1997). Informe sobre índices de Desarrollo Humano en El Salvador. DIGESTYC.
- Programa de las Naciones Unidas para el Desarrollo [PNUD]. (2003). Informe sobre índices de Desarrollo Humano en El Salvador. DIGESTYC.
- Programa de las Naciones Unidas para el Desarrollo [PNUD]. (2013). *Informe sobre sobre Desarrollo Humano 2013*. El ascenso del Sur: Progreso humano en un mundo diverso.
- Programa de las Naciones Unidas para el Desarrollo [PNUD]. (2013-2014). *Informe Regional de Desarrollo Humano 2013-2014 Seguridad Ciudadana con rostro humano: diagnóstico y propuestas para América Latina*.
- Robbins, S. (2015). *Administración*. (12th ed.). Pearson Educación.
<https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnxmY3B5c3RhYWVhcmV6fGd4OjI5NTM0NDQwNjE0ODI4MzE>
- Rodríguez, F. (2004). La pobreza como un proceso de violencia estructural. *Revista de Ciencias Sociales*, 10(001), 42–50.
- Salama, P. (2012). Homicidios, ¿es ineluctable la violencia en América Latina? *Frontera norte*, 25(49), 7-27.
- Santander, A. & Ruiz, R. (2004). *Relación entre variables cuantitativas*. Editorial ECIMED.
- Sauto, R. (2005). *Manual de metodología. Construcción del marco teórico, formulación de los objetivos y elección de la metodología*. CLACSO.
- Schwarz, G. (1978). Estimación de la dimensión de un modelo. *Anales de Estadística*, 6, 461-464. <http://dx.doi.org/10.1214/aos/1176344136>
- Sokal, R. (2002). *Introducción a la bioestadística*. DialnetAnálisisDeVarianzaNoParametrica-6788415.pdf

- Supo, J. (2014). *Metodología de la investigación: Diseño y ejecución*. Universidad César Vallejo.
- Urbano, C. & Yuni, J. (2014). *Recursos metodológicos para la preparación de proyectos de investigación*. Editorial Brujas.
- Vidal, F. (2008). *Los nuevos aceleradores de la violencia remodernizada*. Universidad Pontificia Comillas de Madrid.
- Valiente, A. (2019). Una nueva aproximación estadística al problema de homicidios en El Salvador. *América latina hoy: Revista de ciencias sociales*, 82, 149-173.

Anexo 1. Base de datos

Año	Tasa de Homicidios	Índice de GINI	Desempleo	PIB per cápita	IDH	Pobreza	Gobierno de la Ley
1990	70.00	0.54	7.53	914.13	0.54	59.70	22.10
1991	90.00	0.54	7.53	983.18	0.53	59.70	22.10
1992	120.00	0.52	7.88	1073.31	0.55	58.80	22.10
1993	135.00	0.50	9.94	1216.70	0.56	57.80	22.10
1994	137.96	0.50	7.68	1380.71	0.57	54.50	22.10
1995	141.72	0.50	7.65	1585.11	0.58	46.30	22.10
1996	119.37	0.51	7.68	1684.78	0.59	51.70	22.10
1997	114.39	0.51	7.98	1778.84	0.60	48.33	25.60
1998	96.31	0.54	7.10	1886.36	0.60	45.10	29.00
1999	65.78	0.52	6.68	1930.63	0.61	41.70	28.10
2000	60.31	0.52	6.96	2001.54	0.62	38.80	27.20
2001	60.57	0.51	6.96	2072.30	0.62	38.80	31.90
2002	47.55	0.52	5.73	2124.10	0.63	36.80	36.60
2003	55.99	0.51	6.26	2209.50	0.64	36.10	39.60
2004	64.69	0.48	6.05	2278.43	0.64	34.80	40.20
2005	64.14	0.49	7.22	2428.57	0.65	35.20	38.30
2006	64.60	0.46	6.57	2631.82	0.66	30.90	32.10
2007	57.27	0.45	6.41	2786.16	0.66	34.60	31.10
2008	51.85	0.47	5.88	2933.39	0.66	39.90	30.30
2009	71.16	0.46	7.33	2858.18	0.66	37.80	27.50
2010	64.47	0.44	4.89	2983.23	0.67	36.50	24.20
2011	70.38	0.42	4.30	3266.01	0.67	46.60	29.10
2012	41.58	0.42	3.84	3428.41	0.67	34.50	29.60
2013	40.11	0.43	3.69	3509.53	0.67	29.60	33.80
2014	62.29	0.42	4.15	3589.04	0.67	31.80	36.50
2015	105.23	0.41	4.00	3705.58	0.67	34.90	31.70
2016	83.01	0.40	4.42	3806.00	0.67	32.70	26.90
2017	61.71	0.38	4.39	3910.25	0.67	29.20	22.10
2018	52.02	0.39	4.01	4052.62	0.67	26.30	19.70
2019	36.00	0.39	4.17	4167.73	0.67	22.80	23.60

Anexo 2. Pruebas de los 10 modelos

N°	Modelos	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
1	$\text{Pred(H)}=e^{(3.2658+0.0258 \cdot \text{POBEXT})}$	122.25	122.25	119.44	116.40	106.90	86.52	99.45	91.17	83.88	76.83	71.30
2	$\text{Pred(H)}=e^{(10.2795- 9.1529 \cdot \text{GINI}-0.0007 \cdot \text{PIB})}$	109.62	104.45	117.76	127.91	114.04	98.84	84.11	78.75	55.51	64.62	61.49
3	$\text{Pred(H)}=e^{(3.8440- 1.724 \cdot \text{GINI}+0.0310 \cdot \text{POBEXT})}$	117.18	117.18	117.96	118.37	106.86	82.87	96.30	86.75	74.53	69.42	63.45
4	$\text{Pred(H)}=e^{(4.389 - 3.896 \cdot \text{GINI} +0.126 \cdot \text{DESEM}+0.024 \cdot \text{POBEXT})}$	106.35	106.35	117.58	160.88	111.80	91.48	100.54	96.30	70.96	67.06	64.80
5	$\text{Pred(H)}=e^{(6.710 -2.750 \cdot \text{GINI}-3.160 \cdot \text{IDH}+0.022 \cdot \text{POBEXT})}$	125.46	129.49	134.13	126.09	113.61	91.91	97.56	87.77	75.28	71.51	65.00
6	$\text{Pred(H)}=e^{(5.769 -4.287 \cdot \text{GINI}+0.121 \cdot \text{DESEM}-1.555 \cdot \text{IDH}+0.019 \cdot \text{POBEXT})}$	105.61	107.26	116.19	156.90	110.38	92.66	97.20	93.08	69.20	66.14	63.75
7	$\text{Pred(H)}=e^{(7.619 -5.9583 \cdot \text{GINI}+0.000 \cdot \text{PIB}+0.014 \cdot \text{POBEXR}-0.006 \cdot \text{G\&L})}$	164.80	164.80	183.33	203.66	194.46	173.37	176.17	164.56	128.88	139.20	134.38
8	$\text{Pred(H)}=e^{(9.065 -5.918 \cdot \text{GINI}+0.106 \cdot \text{DESEM}+0.000 \cdot \text{PIB} -3.210 \cdot \text{IDH}-0.004 \cdot \text{G\&L})}$	127.19	131.34	143.89	195.12	148.71	143.55	131.45	129.58	97.50	102.03	102.15
9	$\text{Pred(H)}=e^{(5.677 -4.194 \cdot \text{GINI}+0.121 \cdot \text{DESEM}-1.440 \cdot \text{IDH}+ 0.019 \cdot \text{POBEXR}-0.001 \cdot \text{G\&L})}$	105.42	106.95	115.90	156.40	110.15	92.58	97.31	92.97	69.08	66.04	63.78
10	$\text{Pred(H)}=e^{(5.713 -4.337 \cdot \text{GINI}+0.119 \cdot \text{DESEM}+0.000 \cdot \text{PIB} - 1.256 \cdot \text{IDH}+0.019 \cdot \text{POBEXR}-0.001 \cdot \text{G\&L})}$	110.06	111.45	121.48	164.04	116.27	97.90	102.95	98.48	72.98	70.16	67.85

N°	Modelos	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
1	$\text{Pred(H)}=e^{(3.2658+0.0258 \cdot \text{POBEXT})}$	71.30	67.71	66.50	64.31	64.97	58.15	63.97	73.35	69.48	67.19
2	$\text{Pred(H)}=e^{(10.2795- 9.1529 \cdot \text{GINI}-0.0007 \cdot \text{PIB})}$	64.13	56.44	58.26	73.05	60.01	68.50	67.38	50.62	58.47	64.33
3	$\text{Pred(H)}=e^{(3.8440- 1.724 \cdot \text{GINI}+0.0310 \cdot \text{POBEXT})}$	64.56	59.64	59.37	60.06	59.77	55.08	62.85	71.57	68.22	67.83
4	$\text{Pred(H)}=e^{(4.389 - 3.896 \cdot \text{GINI} +0.126 \cdot \text{DESEM}+0.024 \cdot \text{POBEXT})}$	67.37	52.89	57.81	61.34	69.03	64.47	71.80	70.55	83.73	64.52
5	$\text{Pred(H)}=e^{(6.710 -2.750 \cdot \text{GINI}-3.160 \cdot \text{IDH}+0.022 \cdot \text{POBEXT})}$	66.81	60.27	59.10	62.38	59.32	56.78	63.31	67.34	66.09	65.75
6	$\text{Pred(H)}=e^{(5.769 -4.287 \cdot \text{GINI}+0.121 \cdot \text{DESEM}-1.555 \cdot \text{IDH}+0.019 \cdot \text{POBEXT})}$	66.55	52.07	56.31	60.90	66.69	63.61	69.86	66.51	79.50	61.93
7	$\text{Pred(H)}=e^{(7.619 -5.9583 \cdot \text{GINI}+0.000 \cdot \text{PIB}+0.014 \cdot \text{POBEXR}-0.006 \cdot \text{G\&L})}$	138.67	123.51	127.49	149.16	142.94	167.03	187.84	180.44	189.12	213.40
8	$\text{Pred(H)}=e^{(9.065 -5.918 \cdot \text{GINI}+0.106 \cdot \text{DESEM}+0.000 \cdot \text{PIB} -3.210 \cdot \text{IDH}-0.004 \cdot \text{G\&L})}$	106.35	83.62	89.80	104.63	108.94	120.56	126.26	106.38	133.10	113.51
9	$\text{Pred(H)}=e^{(5.677 -4.194 \cdot \text{GINI}+0.121 \cdot \text{DESEM}-1.440 \cdot \text{IDH}+ 0.019 \cdot \text{POBEXR}-0.001 \cdot \text{G\&L})}$	66.20	51.67	55.71	60.06	66.03	63.26	69.49	66.33	79.44	62.04
10	$\text{Pred(H)}=e^{(5.713 -4.337 \cdot \text{GINI}+0.119 \cdot \text{DESEM}+0.000 \cdot \text{PIB} - 1.256 \cdot \text{IDH}+0.019 \cdot \text{POBEXR}-0.001 \cdot \text{G\&L})}$	70.52	55.20	59.65	64.61	70.89	68.43	75.30	71.75	85.80	67.65

N°	Modelos	2011	2012	2013	2014	2015	2016	2017	2018	2019
1	$\text{Pred(H)}=e^{(3.2658+0.0258 \cdot \text{POBEXT})}$	87.19	63.81	62.67	59.52	64.47	60.91	55.65	51.64	47.18
2	$\text{Pred(H)}=e^{(10.2795- 9.1529 \cdot \text{GINI}-0.0007 \cdot \text{PIB})}$	63.38	56.57	48.77	50.55	51.06	52.15	58.22	48.09	44.37
3	$\text{Pred(H)}=e^{(3.8440- 1.724 \cdot \text{GINI}+0.0310 \cdot \text{POBEXT})}$	96.02	65.98	55.72	60.69	67.97	64.59	59.98	53.89	48.35
4	$\text{Pred(H)}=e^{(4.389 - 3.896 \cdot \text{GINI} +0.126 \cdot \text{DESEM}+0.024 \cdot \text{POBEXT})}$	82.51	58.24	48.86	56.76	62.38	64.87	64.23	54.93	51.53
5	$\text{Pred(H)}=e^{(6.710 -2.750 \cdot \text{GINI}-3.160 \cdot \text{IDH}+0.022 \cdot \text{POBEXT})}$	86.75	66.47	58.06	62.64	68.93	67.50	66.04	60.27	55.81
6	$\text{Pred(H)}=e^{(5.769 -4.287 \cdot \text{GINI}+0.121 \cdot \text{DESEM}-1.555 \cdot \text{IDH}+0.019 \cdot \text{POBEXT})}$	76.12	57.21	49.04	56.43	61.35	64.62	65.63	56.83	54.22
7	$\text{Pred(H)}=e^{(7.619 -5.9583 \cdot \text{GINI}+0.000 \cdot \text{PIB}+0.014 \cdot \text{POBEXR}-0.006 \cdot \text{G\&L})}$	268.89	226.31	194.13	209.08	238.53	252.67	278.95	256.02	238.14
8	$\text{Pred(H)}=e^{(9.065 -5.918 \cdot \text{GINI}+0.106 \cdot \text{DESEM}+0.000 \cdot \text{PIB} -3.210 \cdot \text{IDH}-0.004 \cdot \text{G\&L})}$	117.70	111.87	102.05	112.46	119.72	135.37	154.84	141.53	141.73
9	$\text{Pred(H)}=e^{(5.677 -4.194 \cdot \text{GINI}+0.121 \cdot \text{DESEM}-1.440 \cdot \text{IDH}+ 0.019 \cdot \text{POBEXR}-0.001 \cdot \text{G\&L})}$	75.74	56.90	48.61	55.73	60.83	64.32	65.52	56.93	54.09
10	$\text{Pred(H)}=e^{(5.713 -4.337 \cdot \text{GINI}+0.119 \cdot \text{DESEM}+0.000 \cdot \text{PIB} - 1.256 \cdot \text{IDH}+0.019 \cdot \text{POBEXR}-0.001 \cdot \text{G\&L})}$	82.93	62.35	53.21	61.04	66.73	70.60	72.13	62.63	59.49

Anexo 3. Cálculo de Error Cuadrático Medio

Modelo 1: $\text{Pred}(H)=e^{(3.2658+0.0258 \cdot \text{POBEXT})}$			
Año	Pred(H)	Real(H)	Error
1990	122.25	110	150.04
1991	122.25	105	297.53
1992	119.44	120	0.31
1993	116.40	130	184.94
1994	106.90	115	65.60
1995	86.52	100	181.79
1996	99.45	85	208.82
1997	91.17	80	124.74
1998	83.88	56	777.26
1999	76.83	65	140.07
2000	71.30	60	127.60
2001	71.30	63	68.82
2002	67.71	55	161.55
2003	66.50	57	90.22
2004	64.31	72	59.21
2005	64.97	61	15.78
2006	58.15	69	117.73
2007	63.97	68	16.21
2008	73.35	51	499.45
2009	69.48	57	155.75
2010	67.19	65	4.79
2011	87.19	62	634.49
2012	63.81	55	77.60
2013	62.67	49	186.80
2014	59.52	51	72.52
2015	64.47	52	155.53
2016	60.91	51	98.28
2017	55.65	57	1.81
2018	51.64	49	6.98
2019	47.18	45	4.77
Suma			4687.00
Promedio			156.23

Modelo 2: $\text{Pred}(H)=e^{(10.2795- 9.1529 \cdot \text{GINI}-0.0007 \cdot \text{PIB})}$			
Año	Pred(H)	Real(H)	Error
1990	109.62	110	0.14
1991	104.45	105	0.30
1992	117.76	120	5.00
1993	127.91	130	4.35
1994	114.04	115	0.92
1995	98.84	100	1.35
1996	84.11	85	0.78
1997	78.75	80	1.55
1998	55.51	56	0.24
1999	64.62	65	0.14
2000	61.49	60	2.23
2001	64.13	63	1.28
2002	56.44	55	2.06
2003	58.26	57	1.58
2004	73.05	72	1.11
2005	60.01	61	0.97
2006	68.50	69	0.25
2007	67.38	68	0.38
2008	50.62	51	0.15
2009	58.47	57	2.15
2010	64.33	65	0.45
2011	63.38	62	1.89
2012	56.57	55	2.45
2013	48.77	49	0.05
2014	50.55	51	0.20
2015	51.06	52	0.89
2016	52.15	51	1.32
2017	58.22	57	1.49
2018	48.09	49	0.83
2019	44.37	45	0.40
Suma			36.95
Promedio			1.23

Modelo 3: $\text{Pred(H)} = e^{(3.8440 - 1.724 \cdot \text{GINI} + 0.0310 \cdot \text{POBEXT})}$			
Año	Pred(H)	Real(H)	Error
1990	117.18	110	51.60
1991	117.18	105	148.43
1992	117.96	120	4.17
1993	118.37	130	135.30
1994	106.86	115	66.29
1995	82.87	100	293.36
1996	96.30	85	127.67
1997	86.75	80	45.52
1998	74.53	56	343.19
1999	69.42	65	19.56
2000	63.45	60	11.93
2001	64.56	63	2.43
2002	59.64	55	21.53
2003	59.37	57	5.64
2004	60.06	72	142.66
2005	59.77	61	1.52
2006	55.08	69	193.65
2007	62.85	68	26.49
2008	71.57	51	422.96
2009	68.22	57	125.92
2010	67.83	65	7.98
2011	96.02	62	1157.12
2012	65.98	55	120.66
2013	55.72	49	45.11
2014	60.69	51	93.83
2015	67.97	52	255.03
2016	64.59	51	184.76
2017	59.98	57	8.91
2018	53.89	49	23.91
2019	48.35	45	11.21
Suma			4098.35
Promedio			136.61

Modelo 4: $\text{Pred(H)} = e^{(4.389 - 3.896 \cdot \text{GINI} + 0.126 \cdot \text{DESEM} + 0.024 \cdot \text{POBEXT})}$			
Año	Pred(H)	Real(H)	Error
1990	106.35	110	13.32
1991	106.35	105	1.82
1992	117.58	120	5.83
1993	118.37	130	135.30
1994	111.80	115	10.26
1995	91.48	100	72.62
1996	100.54	85	241.38
1997	96.30	80	265.60
1998	70.96	56	223.89
1999	67.06	65	4.24
2000	64.80	60	23.00
2001	67.37	63	19.10
2002	52.89	55	4.44
2003	57.81	57	0.66
2004	61.34	72	113.65
2005	69.03	61	64.41
2006	64.47	69	20.49
2007	71.80	68	14.43
2008	70.55	51	382.31
2009	83.73	57	714.59
2010	64.52	65	0.24
2011	82.51	62	420.53
2012	58.24	55	10.48
2013	48.86	49	0.02
2014	56.76	51	33.14
2015	62.38	52	107.74
2016	64.87	51	192.26
2017	64.23	57	52.26
2018	54.93	49	35.14
2019	51.53	45	42.65
Suma			3225.80
Promedio			107.53

Modelo 5: Pred(H)=e^(6.710 -2.750 · GINI-3.160 · IDH+0.022 · POBEXT)			
Año	Pred(H)	Real(H)	Error
1990	125.46	110	239.06
1991	129.49	105	599.74
1992	134.13	120	199.62
1993	126.09	130	15.28
1994	113.61	115	1.92
1995	91.91	100	65.46
1996	97.56	85	157.83
1997	87.77	80	60.43
1998	75.28	56	371.68
1999	71.51	65	42.35
2000	65.00	60	25.01
2001	66.81	63	14.54
2002	60.27	55	27.75
2003	59.10	57	4.43
2004	62.38	72	92.60
2005	59.32	61	2.83
2006	56.78	69	149.31
2007	63.31	68	21.96
2008	67.34	51	266.88
2009	66.09	57	82.61
2010	65.75	65	0.56
2011	86.75	62	612.43
2012	66.47	55	131.64
2013	58.06	49	82.11
2014	62.64	51	135.48
2015	68.93	52	286.64
2016	67.50	51	272.41
2017	66.04	57	81.65
2018	60.27	49	127.10
2019	55.81	45	116.79
		Suma	4288.10
		Promedio	142.94

Modelo 6: $\text{Pred(H)} = e^{(5.769 - 4.287 \cdot \text{GINI} + 0.121 \cdot \text{DESEM} - 1.555 \cdot \text{IDH} + 0.019 \cdot \text{POBEXT})}$			
Año	Pred(H)	Real(H)	Error
1990	105.61	110	19.27
1991	107.26	105	5.13
1992	116.19	120	14.55
1993	156.90	130	723.87
1994	110.38	115	21.34
1995	92.66	100	53.86
1996	97.20	85	148.81
1997	93.08	80	171.14
1998	69.20	56	174.27
1999	66.14	65	1.30
2000	63.75	60	14.09
2001	66.55	63	12.57
2002	52.07	55	8.57
2003	56.31	57	0.48
2004	60.90	72	123.13
2005	66.69	61	32.36
2006	63.61	69	29.07
2007	69.86	68	3.48
2008	66.51	51	240.64
2009	79.50	57	506.44
2010	61.93	65	9.40
2011	76.12	62	199.37
2012	57.21	55	4.89
2013	49.04	49	0.00
2014	56.43	51	29.46
2015	61.35	52	87.41
2016	64.62	51	185.41
2017	65.63	57	74.53
2018	56.83	49	61.37
2019	54.22	45	84.95
Suma			3041.17
Promedio			101.37

Modelo 7: $\text{Pred(H)} = e^{(7.619 - 5.9583 \cdot \text{GINI} + 0.000 \cdot \text{PIB} + 0.014 \cdot \text{POBEXR} - 0.006 \cdot \text{G\&L})}$			
Año	Pred(H)	Real(H)	Error
1990	164.80	110	3002.78
1991	164.80	105	3575.75
1992	183.33	120	4010.60
1993	203.66	130	5425.74
1994	194.46	115	6314.62
1995	173.37	100	5383.72
1996	176.17	85	8312.53
1997	164.56	80	7150.72
1998	128.88	56	5312.18
1999	139.20	65	5505.07
2000	134.38	60	5532.63
2001	138.67	63	5725.32
2002	123.51	55	4692.97
2003	127.49	57	4969.36
2004	149.16	72	5953.58
2005	142.94	61	6714.35
2006	167.03	69	9610.16
2007	187.84	68	14360.48
2008	180.44	51	16754.75
2009	189.12	57	17455.71
2010	213.40	65	22021.27
2011	268.89	62	42804.77
2012	226.31	55	29347.76
2013	194.13	49	21062.88
2014	209.08	51	24989.09
2015	238.53	52	34793.61
2016	252.67	51	40670.35
2017	278.95	57	49263.47
2018	256.02	49	42858.07
2019	238.14	45	37303.45
Suma			490877.73
Promedio			16362.59

Modelo 8: $\text{Pred(H)} = e^{(9.065 - 5.918 \cdot \text{GINI} + 0.106 \cdot \text{DESEM} + 0.000 \cdot \text{PIB} - 3.210 \cdot \text{IDH} - 0.004 \cdot \text{G\&L})}$			
Año	Pred(H)	Real(H)	Error
1990	127.19	110	295.40
1991	131.34	105	693.59
1992	143.89	120	570.53
1993	195.12	130	4241.26
1994	148.71	115	1136.18
1995	143.55	100	1896.81
1996	131.45	85	2157.28
1997	129.58	80	2458.20
1998	97.50	56	1722.47
1999	102.03	65	1370.87
2000	102.15	60	1776.26
2001	106.35	63	1879.63
2002	83.62	55	819.28
2003	89.80	57	1075.81
2004	104.63	72	1064.92
2005	108.94	61	2298.27
2006	120.56	69	2658.34
2007	126.26	68	3394.47
2008	106.38	51	3066.97
2009	133.10	57	5791.20
2010	113.51	65	2353.54
2011	117.70	62	3102.50
2012	111.87	55	3234.74
2013	102.05	49	2814.70
2014	112.46	51	3777.92
2015	119.72	52	4585.42
2016	135.37	51	7118.48
2017	154.84	57	9572.90
2018	141.53	49	8562.54
2019	141.73	45	9356.04
Suma			94846.53
Promedio			3161.55

Modelo 9: $\text{Pred(H)} = e^{(5.677 - 4.194 \cdot \text{GINI} + 0.121 \cdot \text{DESEM} - 1.440 \cdot \text{IDH} + 0.019 \cdot \text{POBEXR} - 0.001 \cdot \text{G\&L})}$			
Año	Pred(H)	Real(H)	Error
1990	105.42	110	20.96
1991	106.95	105	3.81
1992	115.90	120	16.84
1993	156.40	130	697.15
1994	110.15	115	23.48
1995	92.58	100	55.09
1996	97.31	85	151.63
1997	92.97	80	168.30
1998	69.08	56	171.04
1999	66.04	65	1.07
2000	63.78	60	14.31
2001	66.20	63	10.26
2002	51.67	55	11.10
2003	55.71	57	1.66
2004	60.06	72	142.58
2005	66.03	61	25.27
2006	63.26	69	32.89
2007	69.49	68	2.23
2008	66.33	51	235.13
2009	79.44	57	503.51
2010	62.04	65	8.74
2011	75.74	62	188.82
2012	56.90	55	3.60
2013	48.61	49	0.15
2014	55.73	51	22.40
2015	60.83	52	77.94
2016	64.32	51	177.34
2017	65.52	57	72.61
2018	56.93	49	62.82
2019	54.09	45	82.69
Suma			2985.40
Promedio			99.51

Modelo 10: $\text{Pred(H)} = e^{(5.713 - 4.337 \cdot \text{GINI} + 0.119 \cdot \text{DESEM} + 0.000 \cdot \text{PIB} - 1.256 \cdot \text{IDH} + 0.019 \cdot \text{POBEXR} - 0.001 \cdot \text{G\&L})}$			
Año	Pred(H)	Real(H)	Error
1990	110.06	110	0.00
1991	111.45	105	41.65
1992	121.48	120	2.20
1993	164.04	130	1158.61
1994	116.27	115	1.61
1995	97.90	100	4.40
1996	102.95	85	322.08
1997	98.48	80	341.39
1998	72.98	56	288.41
1999	70.16	65	26.58
2000	67.85	60	61.60
2001	70.52	63	56.61
2002	55.20	55	0.04
2003	59.65	57	7.03
2004	64.61	72	54.63
2005	70.89	61	97.84
2006	68.43	69	0.32
2007	75.30	68	53.29
2008	71.75	51	430.51
2009	85.80	57	829.30
2010	67.65	65	7.04
2011	82.93	62	437.87
2012	62.35	55	54.06
2013	53.21	49	17.74
2014	61.04	51	100.75
2015	66.73	52	217.08
2016	70.60	51	384.24
2017	72.13	57	229.05
2018	62.63	49	185.77
2019	59.49	45	210.08
Suma			5621.79
Promedio			187.39