



UNIVERSIDAD TECNOLÓGICA ECOTEC

TÍTULO DEL TRABAJO:

ESTIMACIÓN DE LA PÉRDIDA ESPERADA POR PAGO DE SINIESTRO DE
VEHÍCULOS RESPECTO A DAÑOS DIRECTO POR CHOQUE DENTRO DEL
PERIODO 2021-2022 EN LA EMPRESA NEW LIFE ECUADOR

LÍNEA DE INVESTIGACIÓN:

DESARROLLO EMPRESARIAL E INNOVACIÓN

MODALIDAD DE TITULACIÓN

EXAMEN COMPLEXIVO

CARRERA:

ECONOMÍA

AUTOR:

MARÍA PAULA BAZURTO HINCAPIÉ

TUTOR:

MGTR. ELENA BEATRIZ PIEDRA

GUAYAQUIL – 2023

CERTIFICADO DEL PORCENTAJE DE COINCIDENCIAS

Habiendo sido nombrado Elena Beatriz Piedra Bonilla, tutor del trabajo de titulación **ESTIMACIÓN DE LA PÉRDIDA ESPERADA POR PAGO DE SINIESTRO DE VEHÍCULOS RESPECTO A DAÑOS DIRECTO POR CHOQUE DENTRO DEL PERIODO 2021-2022 EN LA EMPRESA NEW LIFE ECUADOR**, elaborado por María Paula Bazurto Hincapié, con mi respectiva supervisión como requerimiento parcial para la obtención del título de Licenciada en Economía.

Se informa que el mismo ha resultado tener un porcentaje de coincidencias del 3 %, mismo que se puede verificar en el print de pantalla a continuación:



FL PA

FIRMA DEL TUTOR
Elena Beatriz Piedra Bonilla



ANEXO N° 14

CERTIFICADO DE APROBACIÓN DEL TUTOR PARA LA PRESENTACIÓN A REVISIÓN DEL TRABAJO DE TITULACIÓN

Samborondón, 10 de Julio del 2023

Magister
Mónica Armijos Santos
Coordinadora Académica
Facultad de Ciencias Económicas y Empresariales
Universidad Tecnológica ECOTEC

De mis consideraciones:

Por medio de la presente comunico a usted que el trabajo de titulación TITULADO: ESTIMACIÓN DE LA PÉRDIDA ESPERADA POR PAGO DE SINIESTRO DE VEHÍCULOS RESPECTO A DAÑOS DIRECTO POR CHOQUE DENTRO DEL PERIODO 2021-2022 EN LA EMPRESA NEW LIFE ECUADOR según su modalidad EXAMEN COMPLEXIVO; fue revisado, siendo su contenido original en su totalidad, así como el cumplimiento de los requerimientos establecidos en la guía para la elaboración del trabajo de titulación, Por lo que se autoriza a: **Maria Paula Bazurto Hincapié** para que proceda a su presentación para la revisión de los miembros del tribunal de sustentación.

ATENTAMENTE,

Elena Piedra Bonilla, PhD

Tutor(a)



ANEXO N°16

**CERTIFICADO DE APROBACIÓN DEL TUTOR PARA LA PRESENTACIÓN DEL
TRABAJO DE TITULACIÓN CON INCORPORACIÓN DE LAS OBSERVACIONES
DE LOS MIEMBROS DEL TRIBUNAL**

Samborondón, 17 de agosto del 2023

Magíster
Ana María Gallardo
Decana de la Facultad
Ciencias Económicas y Empresariales
Universidad Tecnológica ECOTEC

De mis consideraciones:

Por medio de la presente comunico a usted que el trabajo de titulación TITULADO: ESTIMACIÓN DE LA PÉRDIDA ESPERADA POR PAGO DE SINIESTRO DE VEHÍCULOS RESPECTO A DAÑOS DIRECTO POR CHOQUE DENTRO DEL PERIODO 2021-2022 EN LA EMPRESA NEW LIFE ECUADOR según su modalidad de EXAMEN COMPLEXIVO, fue revisado y se deja constancia que el estudiante acogió e incorporó todas las observaciones realizadas por los miembros del tribunal de sustentación por lo que se autoriza a: BAZURTO HINCAPIÉ MARÍA PAULA, para que proceda a la presentación del trabajo de titulación para la revisión de los miembros del tribunal de sustentación y posterior sustentación.

ATENTAMENTE,

PhD. Elena Piedra Bonilla

Tutor(a)

AGRADECIMIENTOS

Agradezco eternamente a mi familia; Marlene, Iván, Marena e Isaac, por su apoyo y fe incondicional hacia mí, por inspirarme a ser mejor en todas las instancias de mi vida, dedico el final de esta etapa a ellos.

Resumen

El presente trabajo sigue los pasos necesarios para la estimación de pérdidas económicas esperadas (PE) por la compañía aseguradora NEW LIFE ECUADOR SA, por concepto de pago de siniestros de vehículos por daño directo por choque, enfocándonos únicamente en siniestros ocurridos durante el periodo 2021-2022 en la ciudad de Guayaquil. Para la estimación, se utilizaron las variables de frecuencia y severidad, basándonos en los datos históricos de la empresa; dentro del análisis también se estima la Pérdida no esperada por el negocio conocida como Valor en Riesgo (VaR) para conocer la probabilidad de la asegurado a la exposición de siniestros atípicos, estas estimaciones se llevaron a cabo mediante la Simulación Montecarlo, ejecutada bajo el enfoque LDA, también conocido Modelo de Distribución de Pérdidas Agregadas. El primer paso es identificar las distribuciones cuyo comportamiento se encuentre bajo los niveles de ajuste conforme al modelo, haciendo uso del criterio Bayesiano, Akaike y Kolmogórov-Smirnov, obteniendo como resultado las distribuciones de Lognormal y Binomial negativa respectivamente, siendo estas las escogidas en base a los niveles de ajuste deseados. Después, se procedió a efectuar la simulación con los datos de la empresa, obteniendo como valor mensual \$ 347.387,00 (PE), por otro lado, el valor de la pérdida no esperada o VaR al 99% dio como resultado \$ 592.614.10. El estudio evidencia la importancia de conocer mediante una aproximación estadística, no solo el valor que se espera asumir en pago de siniestros, sino también la probabilidad de exposición a casos extremos o catastróficos.

PALABRAS CLAVES

Pérdida Esperada, Monte Carlo, Siniestro, Vehículo

Abstract

The present work follows the necessary steps for the estimation of expected economic losses (PE) by the insurance company NEW LIFE ECUADOR SA, for payment of vehicle claims for direct damage due to crash, focusing only on claims that occurred during the period 2021-2022 in the city of Guayaquil. For the estimation, the variables of frequency and severity were used based on the historical data of the company; Within the analysis, the Unexpected Loss for the business, known as Value at Risk (VaR), is also estimated to know the probability of the insured to exposure to atypical claims, these estimates were carried out through the Monte Carlo Simulation, executed under the LDA approach, also known Aggregate Loss Distribution Model. The first step is to identify the distributions whose behavior is under the adjustment levels according to the model, making use of the Bayesian, Akaike, and Kolmogorov-Smirnov criteria, obtaining, as a result, the Lognormal and negative binomial distributions, respectively, these being the ones chosen based on to the desired adjustment levels. Afterward, the simulation was carried out with the company's data, obtaining a monthly value of \$347,387.00 (PE). Conversely, the value of the unexpected loss or VaR at 99% resulted in \$592,614.10. The study shows the importance of knowing, through a statistical approach, not only the value that is expected to assume in payment of claims but also the probability of exposure to extreme or catastrophic cases.

KEY WORDS

Expected Loss, Monte Carlo, Vehicle, Claim

CONTENIDO

Capítulo 1. Introducción	11
1.1. Planteamiento del problema a resolver.....	13
1.2. Justificación	14
1.3. Objetivo general.....	14
1.4. Objetivos específicos	14
Capítulo 2. Marco Teórico	17
2.1. Aseguradora, Funcionalidad y Estructura	17
2.1.1. Antecedentes de los seguros de automóviles.....	17
2.1.2. Pólizas	18
2.2. Problema de asimetría de información en los seguros	19
2.3. Selección Adversa	19
2.4. Pérdidas Esperadas.....	21
2.5. Pérdida Agregada	22
2.6. Valor en riesgo	22
2.7. Revisión de literatura de pérdidas esperadas en seguros de automóviles .	23
Capítulo 3. Metodología del Proceso de Investigación	27
3.1. Enfoque de la investigación	27
3.2. Tipo de Investigación	27
3.3. Período y lugar donde se desarrolla la investigación.....	28
3.4. Universo y muestra de la investigación.....	28
3.5. Metodología Montecarlo	28
3.6. Loss Distributional Approach - LDA	30
3.7. Pruebas de ajuste de modelo	31

3.7.1 BIC y AIC	31
3.7.2. Kolmogórov-Smirnov	31
3.7.3. Q-Q Plot.....	32
3.8. Distribuciones de frecuencia y pérdida.	33
Capítulo 4. Análisis de Resultados	36
4.1. Evolución de casos y pérdidas.....	36
4.2. Estadística Descriptiva.....	37
4.3 Selección de distribución - Severidad	39
4.4. Distribución de Frecuencia	44
4.5. Estimación de Valor operativo de riesgo (Monte Carlo)	48
Capítulo 5. Conclusiones	55
Capítulo 6. Estrategias y Recomendaciones	53
Bibliografía	57
Anexos	63

Capítulo 1. Introducción

La estimación de la pérdida esperada por pago de siniestros de vehículos es un aspecto crítico en la gestión de riesgos de una compañía aseguradora. La empresa NEW LIFE ECUADOR se enfrenta al desafío de desarrollar una propuesta metodológica que permita estimar la pérdida esperada por pago de siniestros de vehículos, específicamente en lo que respecta a daños directos por choque, dentro del periodo 2021-2022.

La necesidad de medir el riesgo está cobrando un gran interés en las entidades financieras, y en concreto en las compañías de seguros. El entorno competitivo en el que se desarrolla la actividad ha impulsado la aplicación de técnicas de evaluación del riesgo (González & Santomil, 2006).

Dentro de la industria ecuatoriana aseguradora, los seguros se clasifican en dos grupos: vida y no vida, de acuerdo con la normativa vigente en el país. (Superintendencia de Compañías Valores y Seguros, 2016). En este estudio, nos enfocaremos en las coberturas relacionadas a pólizas de no vida, específicamente autos, y los riesgos inherentes al uso y circulación de los vehículos de motor, las cuales son obtenidas a partir de la necesidad de los propietarios de autos en poder mitigar costos monetarios frente al caso desafortunado de un accidente, entre estos costos podemos mencionar: responsabilidad civil, costos médicos, lesiones graves, gastos legales, robos o hurtos, y daños directos por choques, este tipo de eventualidades son conocidas como siniestros, por otro lado, se conoce como “reclamo” cuando el asegurado genera una solicitud formal por el evento o accidente ocurrido para obtener la cobertura o indemnización respectiva (Hanafy & Ming, 2021)

La cobertura por daños se divide en dos categorías, pérdidas totales y pérdidas parciales, De acuerdo con las condiciones generales de la póliza de vehículo de la

empresa NEW LIFE ECUADOR. Las pérdidas parciales son aquellas pérdidas cuyo valor de reparación, es decir, el valor de la mano de obra más los repuestos, es inferior al 70% de la suma asegurada actual del vehículo, de lo contrario, serán declaradas pérdidas totales. Este valor de reparación se determina haciendo una inspección realizada por un perito de la aseguradora respecto al siniestro reclamado, quien se encarga de verificar que las averías presentadas en el vehículo guarden relación con el evento, por otro lado, el cálculo de la pérdida resulta del valor de la hora hombre en taller por la mano de obra autorizada, mientras que el valor de los repuestos se dará en base a su cotización según mercado. En este estudio, nos enfocaremos exclusivamente en la estimación de pérdidas ocasionadas por daño directo por choque o colisión, excluyendo de esta manera los pagos por concepto de robo, gastos médicos, o daños a la propiedad, gastos legales, etc.

En relación con la cantidad de siniestros presentados dentro del país, los datos extraídos del Instituto Nacional de Estadísticas y Censo revelaron que, en el primer y segundo trimestre del año 2022, se registraron 10.592 siniestros de tránsito, cifra que representa un incremento del 9% con respecto al 2021 en el mismo periodo. A su vez, entre las clases de siniestros vehiculares de mayor frecuencia encontramos que CHOQUES es la causa del 47% del total de siniestros (INEC, 2022).

Desde el enfoque geográfico, las ciudades de Guayaquil y Quito registraron un 34% del total de las víctimas de siniestros de tránsito a nivel nacional, con 2.275 y 1.151 respectivamente, estas cifras nos permiten evidenciar que la ciudad con mayor presencia de siniestralidad vehicular en Ecuador es Guayaquil (INEC, 2022).

La correcta estimación de la pérdida esperada tiene un rol significativo en la rentabilidad y eficiencia de la empresa NEW LIFE ECUADOR, debido a que ayudan a reducir la incertidumbre de los siniestros futuros (Heba, Cordeirob, Haitham, & Mohamed, 2021), lo que la convierte en un tema de gran importancia para la organización. En este contexto, se presenta una estimación de la pérdida esperada

por pago de siniestro de vehículos respecto a daños ocasionados por choque dentro del periodo 2021 – 2022.

1.1. Planteamiento del problema a resolver

El caso práctico de este trabajo se enfoca en calcular la estimación de pérdida esperada para la provisión de siniestros ocurridos en base a los datos históricos de la empresa NEW LIFE ECUADOR en los años 2021 y 2022, con el fin de ofrecer una metodología para el cálculo de pérdida esperada dentro la cobertura de daños directo por choque, la cual dentro de su clasificación es aquella que más casos genera entre las coberturas del Ramo Auto; dentro de los datos históricos de la empresa de estudio, se evidencia que alrededor del 90% de los casos atendidos corresponden a siniestros por colisión, siendo esta la causal primaria de pérdida en el área de negocio.

Existe poca literatura sobre este tema que es de importancia para las empresas aseguradoras, puesto que dentro de la investigación preliminar se hallaron pocos trabajos similares al propuesto, tanto con respecto a valor de siniestralidad esperada como sobre la utilización del método Montecarlo en su estimación. Sin embargo, existe una investigación previa que buscó modelar la totalidad de los eventos en una ventana de tiempo para Colombia (Medina, Restrepo, & Bedoya, 2015) llevado a cabo bajo un enfoque general del mercado y no de las características propias de una empresa. Por otro lado, Azaare et al. (2022) utilizó datos correspondientes al mercado ghanés de seguro de vehículos para la estimación de la pérdida esperada a través de métodos estocásticos dentro de una empresa, sin embargo, esta investigación abarcó todos los tipos de siniestros vehiculares, a diferencia del presente trabajo cuya novedad es estimar en base a una cobertura particular. En este contexto presentado, surge la siguiente pregunta problemática: ¿Cuál es la probabilidad para pérdidas esperadas asociadas a siniestros por choque de vehículos en la ciudad de Guayaquil en el periodo 2021-2022?

1.2. Justificación

Conocer la pérdida esperada nos permite tomar decisiones financieras informadas que minimicen los riesgos asociados al pago de siniestros, por otro lado, esta estimación facilita a las aseguradoras evaluar la rentabilidad de sus productos y cartera. Si la pérdida esperada es mayor que los ingresos esperados, la compañía puede optar por no ofrecer un determinado producto o ajustar su modelo de negocio. (Shi, Feng, & Ivantsova, 2015).

Además, dichas estimaciones pueden ayudar a las empresas aseguradoras a predecir y planificar pérdidas futuras en caso de que acontezcan eventos imprevistos, como desastres naturales o eventos catastróficos, y cómo dichas variables claves pueden afectar la estabilidad del negocio (Torri, Radi, & Hana, 2022).

1.3. Objetivo general

- Estimar la probabilidad de pérdidas económicas esperadas asociadas al pago de siniestro de daños directo por choque de vehículos en la ciudad de Guayaquil en el periodo 2021-2022

1.4. Objetivos específicos

- Analizar la evolución del pago de siniestros por colisión de la empresa en la ciudad de Guayaquil el periodo 2021-2022
- Identificar las distribuciones de probabilidad asociadas a las series de siniestro de daño directo por choque de vehículo para estimar la pérdida económica mensual esperada, mediante un proceso de simulación Montecarlo.
- Proponer estrategias para reducir pérdidas por siniestros

Este trabajo está dividido en 4 capítulos, además de la presente introducción. En el segundo capítulo se tratará el Marco Teórico, donde se explicará una historia breve sobre las aseguradoras y su funcionamiento, sus objetivos como giro de negocio, así como también los tipos de siniestros o de pérdida, luego en el tercer capítulo, se detallará la metodología del estudio, que estará basada en Montecarlo como modelo de aproximación de pérdida esperada. Se mostrará el análisis de la severidad y frecuencia aplicada a una distribución de pérdida, así como también los criterios de selección de la distribución y las variables a utilizar dentro del estudio del problema, El cuarto capítulo corresponde al análisis de resultados, donde se expondrán los resultados relacionados a la estadística descriptiva, estimaciones del modelo, la selección de la distribución de frecuencia y severidad, los criterios de ajuste utilizados y finalmente los resultados de la simulación de la pérdida agregada del modelo. El capítulo cinco y seis culminan el presente estudio con las conclusiones, estrategias y recomendaciones de la autora respecto a las acciones futuras de la empresa aseguradora NEW LIFE ECUADOR.

Capítulo 2. Marco Teórico

2.1. Aseguradora, Funcionalidad y Estructura

2.1.1. Antecedentes de los seguros de automóviles

Los seguros generales tuvieron sus inicios en el siglo XIV, cuya creación tenía como finalidad compartir el riesgo entre comerciantes que surcaban sus negocios en el mercado marítimo, con el tiempo, este tipo de contratos fueron replicados en otros bienes, como casas, terrenos o incluso mercaderías. No fue hasta 1821 cuando en Francia se empezó a dar un aporte común por parte de dueños de coches de caballo con el fin de solventar algún tipo de accidente o daños a causa de su uso, esquema que a finales del siglo XVIII sería adoptado por un nuevo medio de transporte en Estados Unidos: el vehículo automóvil, con la diferencia de que su función era indemnizar solo daños ocasionados a terceros, y más adelante con la propagación de su fabricación ofrecería las coberturas por daños propios que ahora conocemos (Cevallos & Téran, 2018).

De acuerdo con los datos presentados por el Instituto Nacional de Estadística y Censos INEC (2022) durante el periodo del primer trimestre del 2021 al segundo trimestre de 2022, se pudo evidenciar que la causa principal de accidentes de tránsito fue la imprudencia al conducir, siendo esta razón un 38% del total de los siniestros. A su vez, se registró que a nivel nacional la ciudad donde se registra mayor concentración de lesionados es Guayaquil, ubicándose con un 23.77%, seguida de Quito con un 11,24%. Por otro lado, alrededor del 53% de los siniestros involucraron vehículos livianos, ya sean estos motos, autos, SUV o camionetas, en el periodo de enero del 2021 a junio del 2022 (INEC, 2022)

Acorde a los datos registrados por la Superintendencia de Compañías, Valores y Seguros, en el año 2021 la prima neta emitida en el país fue de aproximadamente 1.741,84 millones de dólares. Al segmentar este valor por ramos de cobertura, podemos detectar que el ramo de auto se encuentra en el puesto número dos,

posicionándose por debajo solamente del ramo de vida colectiva, con una participación del 26,27 %, recaudando un total de 452,88 millones de la prima neta total (Calero, 2022).

2.1.2. Pólizas

Las pólizas son contratos mediante los cuales una de las partes (asegurador), se ve en la obligación de indemnizar a la otra parte (asegurado) a cambios del pago de un valor acordado en la compra del seguro, conocido como prima. Cuando el asegurado reporta un reclamo, se procede a levantar los informes periciales que permitan constatar que dichos daños guarden relación con el evento causante del siniestro, y a la vez, que dicho evento se halle con cobertura en base a lo que indique el contrato adquirido previamente, y que, además, figure dentro de las condiciones particulares o generales de la póliza.

A causa del contenido técnico en los contratos de seguros, las pólizas suelen dividirse en condiciones generales y particulares. Las generales cumplen con el objetivo de ser cláusulas prediseñadas que permiten la negociación preliminar estandarizada y contratación en masa con individuos o agentes distintos, por lo tanto, estas son redactadas con la intención de definir elementos esenciales en la relación comercial y legal de las partes, mientras tanto, las condiciones particulares responden al conjunto de cláusulas aisladas resultantes de lo que los agentes, aseguro y aseguradora, hayan decidido negociar individualmente, esto quiere decir que dos asegurados con características similares se pueden encontrar asegurados en una misma compañía de seguros, pero, a causa de distintas negociaciones, gozar de exclusiones o inclusiones diferentes (Giraldo, 2002).

Estos dos apartados poseen beneficios distintos para la empresa aseguradora, Mientras las condiciones generales permiten la uniformidad para efectos jurídicos y evitan la aplicación de criterios en la interpretación del contrato, las particulares permiten flexibilidad y competitividad para el negocio. Sin embargo, parte de los riesgos que la aseguradora asume a la hora efectuar esta transacción reside en las particulares características y comportamientos que cada individuo posee al

momento de adquirir un contrato de póliza, información que no será disponible hasta que se presente un reclamo, lo que implica que el asegurado tenga más conocimiento del riesgo que el asegurador, a este problema se lo conoce como asimetría de información.

2.2. Problema de asimetría de información en los seguros

Compartir el riesgo incurrido mediante un fondo común nos permite establecer una relación estable entre precios, primas y pagos por siniestros, no obstante, para las compañías de seguros, establecer esta relación puede tornarse complejo al no poseer toda la información disponible acerca de los contratantes que deciden adquirir una póliza, fenómeno conocido como asimetría de información, la cual es una falla de mercado que consiste en que ciertos individuos poseen más conocimiento que otros sobre un escenario similar, o en este caso, una relación contractual, afectando el comportamiento natural de ambos agentes y sus decisiones.

Esta problemática de carácter económico puede comprometer los objetivos del agente que se encuentre en desventaja, para poder mitigar los efectos negativos de la información asimétrica, aunque el hecho que estas sucedan es inevitable en un mercado imperfecto es importante proporcionar los incentivos adecuados para que pueda ser controlada (Stiglitz, 2002).

2.3. Selección Adversa

Dentro del Mercado asegurador, se presenta un modelo de selección adversa cuando el asegurado mantiene información privada a la hora de iniciar la relación comercial, lo cual impide a la contraparte (el asegurador) conocer qué tipo de cliente será dicho individuo en el futuro, ya que no se tiene conocimiento previo de las características únicas del asegurado, como sus horas de sueño, su relación con el alcohol, la imprudencia al manejar, su tendencia a conducir a altas velocidades,

entre otras. Al no ser estos rasgos distinguibles entre un asegurado y otro, afecta a las primas y deducibles que todos los asegurados, buenos o malos, asumen al momento de la contratación de un seguro (Macho-Stadler & Pérez-Castrillo, 2001). En la Figura 1 podemos observar la información limitada con la que cuenta el agente (A) antes de la contratación de una póliza por parte del cliente (N) quién cuenta con el poder de elegir no solo la póliza sino su comportamiento una vez realizado el acuerdo, poniéndose así en una posición de ventaja frente a (A)

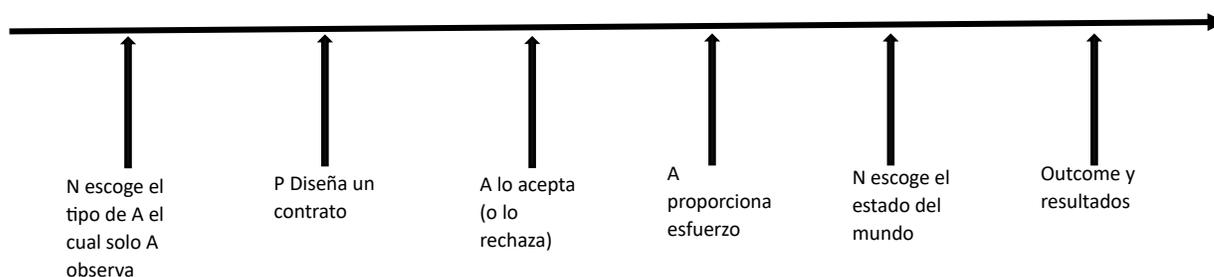


Figura 1. Mecanismo de Selección Adversa

Fuente: Adaptado de Macho-Stadler & Pérez-Castrillo (2001)

Los ramos de seguros suelen ser estudiados en dos categorías, vida y no vida, este último conocido también como generales, dentro de estos existe la rama de vehículos. Los siniestros de auto corresponden a un sistema de compensación de daños que cubren los riesgos resultantes de la tenencia de un vehículo, a cambio de una prima anual o mensual. Estos seguros funcionan como un mecanismo que intenta proteger el orden económico y compartir el costo de las eventualidades a los que todo conductor se encuentra expuesto en la vida diaria, conocidos por el nombre de siniestros, los cuales pueden significar bajos, medios o altos montos de pérdida monetaria para la empresa aseguradora.

En el caso del seguro de automóviles, entre los daños que acceden a cobertura se pueden mencionar los siguientes: la destrucción total del vehículo como consecuencia de un accidente o por daños ocasionados por terceros, esta destrucción es reclamada como total si los repuestos y mano de obra necesarias

para la reparación del vehículo tienen un valor mayor al 70% del valor comercial, por el contrario, si los valores de reparación no superan el 70%, estas serán atendidas como parciales.

Es importante mencionar que dentro del contrato pueden manejarse otras coberturas como son el robo total de vehículos o el robo de partes y accesorios, los cuales se denominan robos parciales, así como también daños a la propiedad pública o privada, muerte, lesiones graves, gastos médicos, gastos legales, lucro cesante, muerte de un tripulante, lesiones graves, entre otras coberturas de la póliza. (Pintado, 2019).

Sin embargo, la existencia de un seguro en sí puede elevar el costo y la probabilidad de ocurrencia de los mismos siniestros, puesto que al existir la garantía de un seguro que cubra los daños de un evento, se debilitan los incentivos para una conducción segura y pueden incentivar a los contratantes de pólizas a no manejar de forma prudente, proporcionar informes falsos, reclamos inflados o incluso actos de fraude, acciones que no pueden ser perfectamente observables y por lo tanto se convierten un problema de información asimétrica para la compañía de seguro

Debido a que no es posible conocer a ciencia cierta el tipo de individuo con quien el asegurador intercambia el riesgo, como respuesta a esta problemática se crea la figura de los deducibles, los cuales las compañías de seguros utilizan para 1) evitar el manejo de reclamos pequeños 2) compartir el riesgo incurrido con el cliente. En base a esta premisa, consideramos el valor de pérdida neta en un evento como el resultado de la pérdida bruta menos el deducible (Transportation, 1999).

2.4. Pérdidas Esperadas

Las pérdidas agregadas pueden definirse como el monto de capital que puede perder una institución como resultado de la exposición al riesgo, representada por la media de la distribución de pérdida, la cual se suele graficar como una línea punteada; por otro lado, la pérdida no esperada hace referencia a la variabilidad de

la distribución de pérdida, la cual se calcula como la diferencia entre el percentil de la distribución menos pérdida esperada, siendo el percentil aquel que se escoja de acuerdo al nivel de confiabilidad, este percentil es conocido como Medida de Pérdida inesperada o Valor en Riesgo (De la Fuente, 2012)

Todo pago de siniestros se deriva de una compensación realizada por la aseguradora a su asegurado cuando este sufre un evento adverso considerado dentro de la cobertura de una póliza contratada, por definición, todo siniestro es un evento imprevisto o accidental, por lo tanto, poder prevenir el pago esperado de los siniestros es el problema central del giro del negocio, esto se conoce como transferencia de riesgo (Magri, Farrugia, Valletta, & Grima, 2019).

Uno de los mayores desafíos para una compañía aseguradora es la posibilidad de que el capital destinado a asumir las pérdidas por siniestros exceda las reservas establecidas a causa de la presencia de siniestros atípicos o catastróficos, que deben ser incluidos dentro de la modelación de la estimación de pérdida, por lo tanto, si este problema no es detectado a tiempo, puede generar a futuro la quiebra real o insolvencia, trayendo consigo aumento de precios, reducción en las ventas y problemas de disponibilidad a nivel del mercado (Cummins, Doherty, & Lo, 2002)

2.5. Pérdida Agregada

Podemos definir distribución de la pérdida agregada como el monto final que la aseguradora deberá asumir durante un determinado periodo de tiempo, esta información se conoce a partir del estudio de datos históricos de la compañía de seguros, es decir, pagos realizados agrupados por tipos de pérdida, con el fin de predecir patrones como el promedio del número de reclamos atendidos, el monto promedio de pago por póliza, el monto esperado de siniestralidad, y el valor del VAR (Septiany, Setiawaty, & Purnaba, 2020).

2.6. Valor en riesgo

El Valor en riesgo (VaR - Value at Risk) es considerado el cuantil de pérdida de valor bajo un período determinado y un nivel de confiabilidad, el cual nos permite calcular el valor monetario a considerar en el caso de ocurrir eventos desfavorables, es decir, estimar la posibilidad de un mal resultado dentro de la operación y cuánto dinero deberá de ocuparse en caso de que dicho evento ocurra, este método utiliza información histórica para simular la distribución de pérdidas que puede sufrir una cartera, por tal, se convierte en una herramienta útil para la gestión de reclamos, debido a que le facilita a la compañía de seguros una Medida clara de la pérdida potencial a la hora de asumir un riesgo (Yousof, Tashkandy, Emam, Ali, & Ibrahim, 2023). Una de las ventajas que representa la obtención del VaR, es que nos permite conocer en un horizonte de tiempo a qué montos de pérdida en el futuro la compañía puede exponerse, considerando la volatilidad que puede generarse por factores de riesgo asociados al mercado (Sotomayor, 2012)

La literatura presenta como alternativa para estimar la pérdida esperada la utilización del método Monte Carlo, el cual funciona bajo un principio de aleatoriedad (Barbu & Zhu, 2020). Consiste en crear la simulación de escenarios de pérdida respecto a eventos correlacionados, considerando la frecuencia y severidad del evento en periodos definidos de tiempo, en este caso, las pérdidas ocasionadas por el choque de vehículos particulares, dando como resultado un número adecuado de escenarios hipotéticos, considerando una muestra de naturaleza homogénea. El método Monte Carlo será descrito a mayor profundidad en el capítulo de Metodología.

2.7. Revisión de literatura de pérdidas esperadas en seguros de automóviles

En la revisión de literatura se buscó referencias ligadas al estudio de pérdidas esperadas en seguros de automóviles o también métodos de análisis de pérdida esperada utilizada en otros ramos de seguros.

En la **Tabla 1**, se puede observar un resumen de las principales referencias utilizadas para la estructuración del presente trabajo, dando énfasis en la metodología y principales resultados. En primer lugar, tenemos un estudio realizado

por Medina, Restrepo y Bedoya (2015), los cuales se centraron en modelar la distribución de severidad de las pérdidas monetarias causadas por robos de vehículos mediante los datos recolectados de varias compañías aseguradoras ubicadas en la ciudad de Bogotá, a diferencia del presente trabajo, se evidencian comportamientos de distribución distintas, donde se analiza la cobertura de robo total de vehículos y no de daño por choque, sin embargo, utiliza las mismas variables de estudio: Severidad y Frecuencia.

El Segundo documento trata sobre un trabajo de tesis realizado por Rendon (2014) donde aborda la pérdida esperada, así como también la máxima pérdida posible con relación a siniestros de Cáncer de Mama, siendo este un ejemplo de la metodología aplicada a ramo de vida y no automotriz como es nuestro caso, de este trabajo se resalta el énfasis en la estimación de pérdidas poco probables y catastróficas mediante el cálculo del VaR, indicadores esenciales para cualquier aseguradora debido a la naturaleza del negocio y la exposición constante al riesgo.

En tercer lugar, se encuentra un documento teórico, Modelo de Distribución de Pérdida Agregada, basándose a diferencia de la otra literatura en los lineamientos de Basilea; de este trabajo se adopta el Periodo de Ajuste de Frecuencia y Severidad a partir del criterio bayesiano.

Como último documento, se presenta una propuesta metodológica a partir de los estudios realizados en la ciudad de Ghana, el cual plantea una metodología de estimación para el cálculo la pérdida esperada y la importancia de éste dentro del mercado asegurador, así como también su rol fundamental en la evaluación del riesgo; resaltando como la poca atención del indicador puede guiar a ineficiencias del mercado. Dentro de la propuesta, a su vez sugiere la utilización de la distribución Lognormal para modelar la severidad de los casos cuando estemos frente a datos de esta naturaleza; como similitud al trabajo Medina, Restrepo y Bedoya (2015), los datos utilizados provienen de distintas compañías aseguradoras.

Tabla 1 – Resumen de revisión de literatura

País	Autor – Título	Método de Estimación	Variable	Unidades de Observación	Puntos Clave
Colombia	(Medina, Restrepo, & Bedoya, 2015)	Basado en el VaR mediante Montecarlo	Monto de siniestros pagados por hurto	Siniestros Ciudad de Bogotá (varias empresas)	De este trabajo el autor logró calcular la distribución severidad que se incurre en Colombia respecto al robo vehicular, tanto para vehículos asegurados como no asegurados
México	(Rendón, 2014)	Distribución de pérdida agregada Monte Carlo	Siniestros reportados	Empresa aseguradora en la ciudad de México	Se establece la máxima pérdida agregada esperada y la mínima, Explica el peor escenario posible en Rangos de pérdida
Colombia	(Macias, Parra, & Carvajal, 2018)	Modelo de Distribución de pérdida agregada	Análisis estadístico y proceso de modelamiento	Documento Teórico	Evaluación datos atípicos, de normalidad y de característica asintótica de extremos
Ghana	(Azaare, y otros, 2022)	Selección de modelos de distribución	Modelo de Selección de criterios de información	Documento Teórico	Modelos adecuados para medir la exposición al riesgo de las aseguradoras de Autos y cuál se aproxima más a la Data

Fuente: Realizado por la autora.

Capítulo 3. Metodología del Proceso de Investigación

3.1. Enfoque de la investigación

El enfoque utilizado dentro de la investigación es cuantitativo, ya que utilizaremos herramientas estadísticas para determinar la probabilidad de que ocurra cierto evento, en este caso, el evento o fenómeno de estudio sería la pérdida esperada producto de siniestros por choque, y su medición permitirá predecir dicho fenómeno en el futuro, basado en la información histórica de los propios siniestros (Sampieri, 2017).

3.2. Tipo de Investigación

En el presente trabajo se utilizará un Enfoque Cuantitativo de naturaleza correlacional, en concreto, se utilizarán los datos de los reclamos pagados por razón de siniestros por choque vehicular en la ciudad de Guayaquil de la empresa NEW LIFE SA para el periodo 2021-2022.

Un alcance correlacional permite conocer el grado de asociación que pueden existir entre dos variables de estudio (Sampieri, 2017), en el caso de esta investigación, la interacción entre la severidad y la frecuencia de los reclamos, permitiendo una predicción de la pérdida esperada agregada. De la combinación de estas dos distribuciones se llevarán a cabo simulaciones utilizando la técnica Montecarlo, obteniendo como resultado la función de pérdidas agregadas y finalmente los valores en riesgos (Value at Risk) junto con el pronóstico de la pérdida esperada.

Existen varias técnicas para estimar la función de pérdida agregada en la literatura, por ejemplo: Monte Carlo, Recursiva, Panjer, y la Transformación de Fast Fourier (FFT). (Shevchenko, 2012). Septiany et al. (2020) utiliza el método Montecarlo para la estimación de la pérdida agregada, de la cual estima los valores en riesgo y el déficit esperado a partir de los reclamos individuales.

3.3. Período y lugar donde se desarrolla la investigación

Los límites temporales donde se desarrollan el presente estudio serán de enero del 2021 a diciembre 2022, tomando todos los siniestros atendidos por la compañía de Seguros NEW LIFE ECUADOR ocurridos en Guayaquil, se decidió no utilizar los datos propios del año 2020 debido a las irregularidades que estos presentan a causa de la pandemia originada por el COVID 2019, puesto que existió una baja significativa de reclamos que podría sesgar los resultados obtenidos dentro del estudio.

3.4. Universo y muestra de la investigación

La población de estudio escogida está conformada únicamente por siniestros de vehículos livianos con pólizas activas que cumplan con las características de daño parcial bajo la definición propia del negocio (la cual fue definido en el Marco Teórico) dentro de la ciudad de Guayaquil, excluyendo de esa manera vehículos pesados, reclamos fuera del área de estudio, los cuales poseen una cuantía mayor, de esta manera se logra proteger la homogeneidad de los datos observados.

3.5. Metodología Montecarlo

La razón por la que se vuelve vigente el análisis de datos y la estadística para la gestión de las empresas de seguros se lo puede atribuir a distintos factores, entre ellos podemos mencionar la importancia de los seguros como actividad económica a nivel mundial

El método Monte Carlo es una herramienta de simulación que tiene como objetivo la estimación, inferencia o simulación de sucesos, que permite estudiar sistemas complejos donde interactúan una variedad de componentes, los cuales forman un ambiente de incertidumbre, mediante la simulación de eventos probabilísticos

(Barbu & Zhu, 2020). De acuerdo con Banks (1999). La simulación es la imitación del funcionamiento de un proceso o sistema del mundo real a lo largo del tiempo, dando como resultado una historia artificial del sistema, permitiendo inferir y analizar mediante su observación el comportamiento del sistema real que se representa.

Para las Compañías de Seguros la naturaleza del giro de negocio es la atención de los eventos de pérdida por naturaleza de siniestros, lo que hace particularmente útiles a las metodologías de riesgo operacional, ya que estas están orientadas a aproximar la severidad de los eventos agregados. En concreto, la presente investigación se desarrollará bajo el enfoque de *Loss Distributional Approach* (LDA), el cual forma parte de la medición de Riesgo Operacional dentro del subconjunto de Métodos de Medición Avanzada, definido por las normativas de Basilea¹ las cuales regulan a los bancos internacionalmente (Macias, Parra, & Carvajal, 2018).

El análisis bajo el método Monte Carlo se realiza mediante la siguiente serie de pasos:

Primero, seleccionamos aleatoriamente una cantidad N de reclamos para la frecuencia, y a su vez seleccionamos N reclamos (X_1, X_2, \dots, X_n) a partir de la distribución de severidad, luego, se calcula la distribución S , que se puede definir como la distribución de probabilidad asociada con la pérdida o la cantidad pagada debido a la pérdida. Posterior a ello, se aplica una suma $(X_1 + X_2 + \dots + X_n)$; repitiendo esta operación reiteradas veces, creando de esa manera las simulaciones (Rendón, Modelación de Pérdida Agregada Aplicada a siniestros de Cáncer de mama: Caso de una empresa aseguradora , 2014).

Los modelos de Pérdida Agregada aplicados al pago de siniestros puede presentar limitaciones al extraer sus datos de información histórica, ya que el número esperado de reclamos recibidos puede variar al incrementarse el número de asegurados de un año a otro, así como también otras variables como la inflación y

¹ El Comité de supervisión Bancaria Basilea es una organización sin fines de lucro encargada de la creación de estándares bancarios internacionales que permiten analizar el riesgo por medio de normativas comunes (Elorriaga, 2010)

la presencia de datos heterogéneos, donde hay valores asegurados y deducibles distintos, esta limitación se puede sobrellevar cuando se plantea el supuesto de que al combinar los datos estamos obteniendo la distribución hipotética del tamaño de la pérdida, construyendo así un modelo más preciso y flexible examinando frecuencia y severidad como variables independientes (Krugman, Panjer, & Wilmot, 2019).

3.6. Loss Distributional Approach - LDA

La metodología *Loss Distributional Approach* basada en la simulación Monte Carlo nos otorga claridad a la hora de evaluar eventos inciertos, partiendo de las distribuciones de frecuencia (cantidad de reclamos) y severidad (tamaño del siniestro) las cuales se eligen en base a información histórica (Medina, Restrepo, & Bedoya, 2015). Es importante mencionar que dichas variables son independientes y deben analizarse por separado, ya que el reporte frecuente de eventos dentro de una misma póliza no debe implicar que la magnitud de los daños sea de gran tamaño, y la poca ocurrencia de reclamos previos en una póliza tampoco garantiza un siniestro de baja cuantía en el futuro. Así mismo, un mayor número de reclamos reportados con respecto a un mes anterior no debe traducirse directamente a una pérdida económica mayor para la compañía de seguros, puesto que esto dependerá del aumento de severidad del daño reclamado (Wang, Qian, & Liu, 2019).

Esta simulación permite la generación de un flujo de numerosos escenarios aleatorios siguiendo la distribución estimada en base al comportamiento histórico de la muestra (Barbu & Zhu, 2020). Para lograr esta distribución se utilizan distintas distribuciones, entre ellas Gamma, Pareto, Exponencial y Lognormal para modelar los reclamos de pérdida, con el fin de encontrar el que se ajuste a los datos en base al criterio bayesiano de información (BIC) (Mazviona & Chiduzza, 2013) y el criterio de Akaike (AIC).

3.7. Pruebas de ajuste de modelo

Se suele asumir que la data sigue una distribución normal, sin embargo, en la práctica no siempre resulta de esta manera, por esta razón dentro de la estadística se han diseñado distintas pruebas de ajuste que permitan evaluar el comportamiento de los datos de estudio, a continuación describen los criterios que se utilizarán para llevar a cabo la selección de la distribución (Aslam, 2020).

3.7.1 BIC y AIC

El criterio Bayesiano de Información (BIC) y el criterio de Akaike (AIC) nos sirve dentro del análisis de datos para seleccionar entre dos o más distribuciones alternativas y cuál de estas se ajusta mejor a las necesidades del estudio en cuestión, basándonos en que tan complejo o sencillos estos se comportan. Normalmente la distribución con un BIC y AIC menor es el escogido. El BIC como criterio como fue desarrollado por Gideon E. Swartz en 1978 a partir de los principios de la estadística bayesiana, este método de selección normalmente tiende a penalizar los modelos de más alta complejidad y ponderar aquellos más parsimoniosos (Pitt & Myung, 2002).

Con el cálculo del BIC, logramos obtener la distribución más probable a futuro cuyo comportamiento sea más parecido o apegado a la distribución de probabilidad del modelo real (Bauldry, 2015). En relación con el criterio de Akaike, desarrollado por Hirotugu Akaike a finales de los años setenta, este escoge el modelo en base su verosimilitud y número de parámetros libres que este posea (Parze, 1988). Ambos modelos se encuentran estrechamente relacionados y diversos autores recomiendan el uso de ambos dentro de la selección de modelos de probabilidad.

3.7.2. Kolmogórov-Smirnov

Se trata de una prueba estadística empleada para determinar la bondad de un ajuste, la cual se calcula a partir de la diferencia mayor entre lo observado y lo teórico, calculando la distancia que hay entre ambos, contraste que permite definir si la distribución que se está escogiendo es la apropiada para el análisis de la muestra, se considera una prueba no paramétrica porque no asume ninguna distribución subyacente en particular (Ríos & Peña, 2020).

El método de Kolmogórov-Smirnov es también conocido como K-S y pertenece a las diversas estadísticas de bondad de ajuste, denominadas Distribuciones Empíricas de función, la cual está basada en la comparación de la población acumulada (Hanusz & Tarasińska, 2015). Las pruebas K-S se encuentran basadas en el supuesto de que el valor p (P Value) sea menor al 5%, lo que indica que los datos sí siguen la debida distribución a la que le estamos intentando asimilar, en caso de que no cumpla, esta se rechaza, esta hipótesis es conocida como hipótesis nula (Oktaviana & Irhamah, 2021)

3.7.3. Q-Q Plot

La herramienta de Q-Q plot, llamada así por sus siglas en inglés *quantile-quantile* es uno de los métodos de diagnóstico más utilizados para demostrar gráficamente el ajuste de la distribución que se está presentando, partiendo de una muestra aleatoria de datos, sin embargo, la diversa literatura recomienda acompañarlo de métodos de ajuste numéricos tales como BIC, AIC, K-S, entre otros, lo cual nos permita generar mayor evidencia o robustez en el modelo para permitir mayor certeza a la hora de efectuar conclusiones (Razal & Bee Wah, 2011).

Los gráficos Q-Q actúan como diagramas de dispersión, que permiten al usuario visualizar líneas de referencia o bandas de confianza, los puntos que se distancian de estas bandas denotarían la falta de ajuste, lo cual facilita de interpretación de cómo se está comportando la distribución evaluada en relación con los datos utilizados en la investigación (Almeida, Loy, & Hofmann, 2018)

3.8. Distribuciones de frecuencia y pérdida.

La variable de Frecuencia hace referencia a la cantidad de eventos que suceden en un tiempo determinado, lo que permite obtener la probabilidad de ocurrencia (Macias, Parra, & Carvajal, 2018), puede esta expresarse como:

$$P_{i,j}(n) = Prob(N_{i,j} = n) \quad (1)$$

Esta variable (P) se considera como una variable aleatoria discreta y permite observar una N cantidad de eventos; diversos autores proponen utilizar distribuciones discretas como Poisson y la Binomial Negativa (Franco & Murillo, 2008)

Por otro, la variable de Severidad busca analizar el impacto o magnitud de las perdidas representadas en una distribución continua, la cual se encuentra basada en datos históricos de pérdidas sufridas en un periodo, para las estimaciones de pérdida agregada se recomienda la distribución Lognomal o la de Weibull (Franco & Murillo, 2008), también pueden utilizarse Gamma, Gumbel o Pareto (Macias, Parra, & Carvajal, 2018), dicha variable puede expresarse matemáticamente como:

$$F_{i,j}(x) = Prob(x_{i,j} = x) \quad (2)$$

Esta Variable, siendo X el vector de severidad y F la distribución teórica

En el siguiente diagrama (Figura 2) podemos ver un paso a paso sobre cómo se llevará a cabo la estimación de pérdida esperada que se realizará para el estudio de los datos, los cuales son obtenidos del reporte mensual de promedios de la compañía en la sucursal de Guayaquil, y es alimentado mes a mes y tabulado para conocer los resultado del cierre del mes y la factura promedio que se está manejando, de esta base de datos se observan dos variables: la fecha del evento, que es igual a la fecha del reporte del asegurado, y el monto de exposición, es decir la suma de todos los pagos del reclamo asumidos por la compañía de seguros, restando el deducible de participación del cliente y sin incluir valores de subrogación o recuperos.

Finalmente, estas variables nos sirven para modelar 1) La distribución de severidad o tamaño del pago de los siniestros y 2) La frecuencia en la que los eventos ocurren; para lograr seleccionar la distribución que más se ajuste, se utilizarán el criterio bayesiano y akaike.

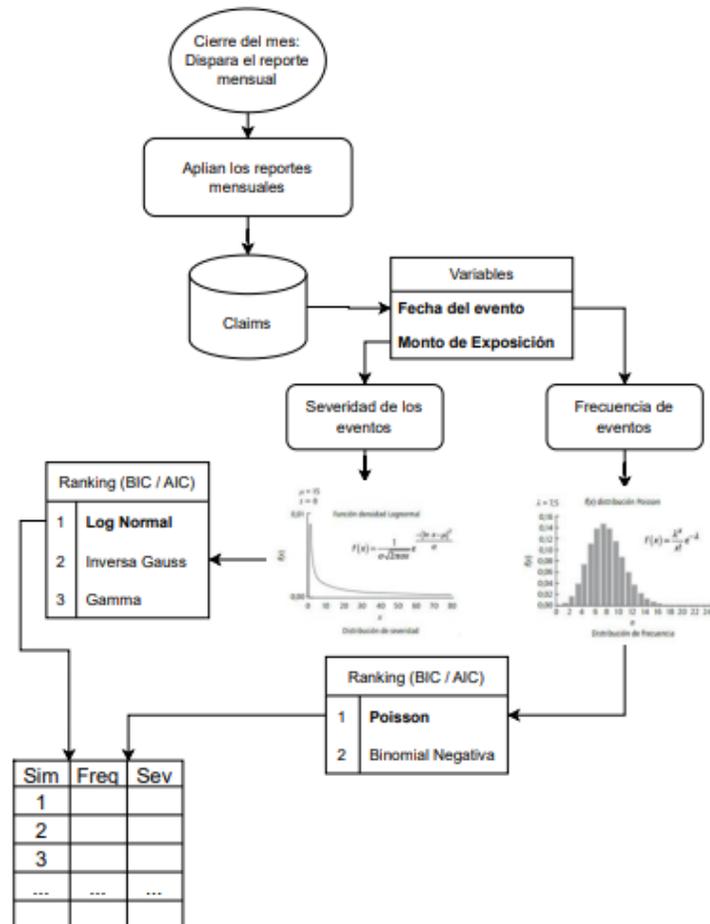


Figura 2. **Flujograma de Proceso**

Fuente: Realizado por la autora

Capítulo 4. Análisis de Resultados

A continuación, se demuestran los resultados obtenidos partiendo de la estimación de la pérdida esperada por siniestros pagados en la cobertura de daño directo por choque de vehículos, los datos corresponden a los reportes mensuales de siniestros reportados en el área de vehículos por la empresa NEW LIFE ECUADOR SA en la ciudad de Guayaquil desde enero del 2021 a diciembre del 2022, y la cuantía de estos.

4.1. Evolución de casos y pérdidas

En la Figura 3 se encuentran presentada la evolución mensualizada de los siniestros por daños directos por choque en la ciudad de Guayaquil, en base a los datos históricos presentados en el 2021 y 2022 de la compañía NEW LIFE ECUADOR, demostrando los rangos de severidad, que se refiere a los montos de pago (izquierda) y frecuencia, que indica la cantidad de reclamos reportados por los asegurados (derecha).

Utilizando los datos obtenidos de cada mes estudiado respecto al pago de siniestros parciales por cobertura de choque; se evidencia que meses como el mes de marzo es donde se registra mayor siniestralidad en comparación a los otros meses, esto probablemente puede estar relacionado al feriado de carnaval, en el cual se reportan más siniestros históricamente, por otro lado, febrero es de baja siniestralidad, siendo este mes donde se reportan menos reclamos. lo cual puede deberse a que se trata de un mes corto, tendencia que se cumple en los dos años donde da lugar el estudio.

Por otro lado, también observamos según los registros de la aseguradora, que en el año 2021 fueron atendidos mas siniestros, y que el pago mensual por siniestro de vehículo se encuentra entre los 300,000.00 y 500,000.00. Además, se aprecia que en el 2021 y 2022 se pagaron 4,8 millones y 3.4 millones de dólares

respectivamente en relación con reclamos de daño de directo por choque, convirtiéndolo en un rubro significativo para la empresa NEW LIFE ECUADOR SA y reafirmando la importancia del análisis dentro del giro del negocio

Frecuencia y severidad agregada observada

Periodo de estudio: 2021 - 2022

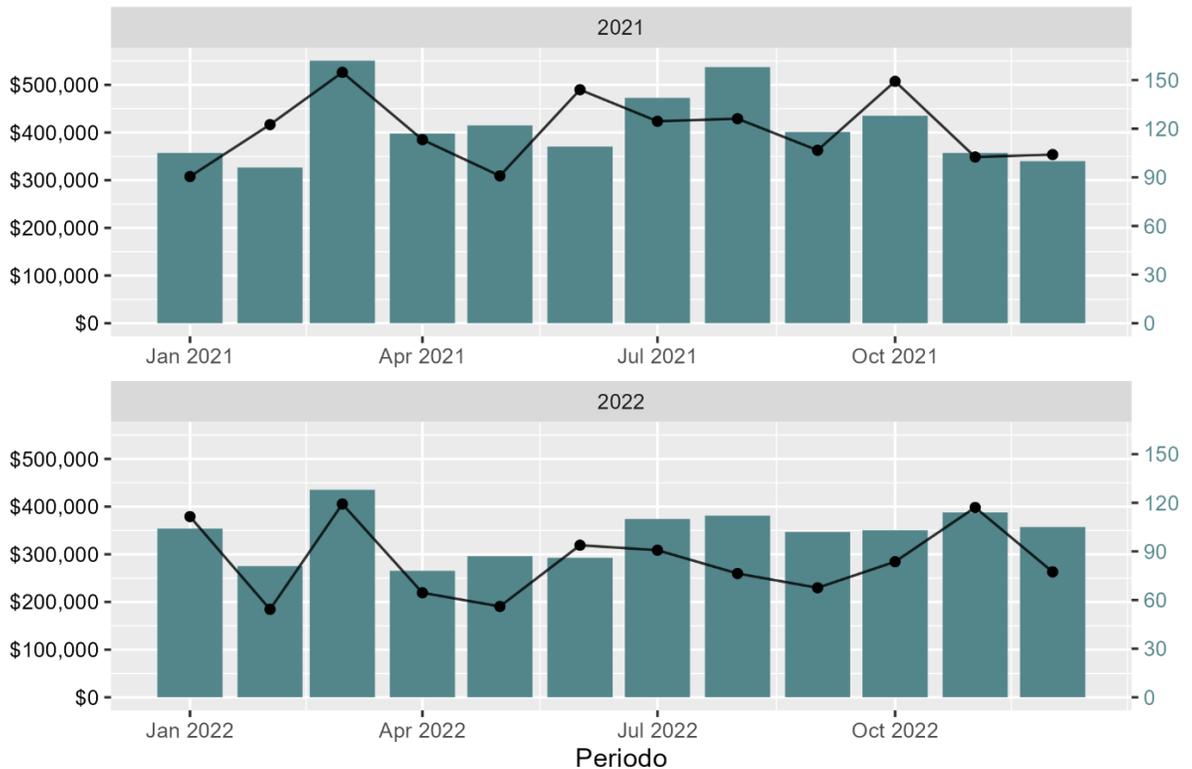


Figura 3.- Evolución de número de casos y valor de pérdida de siniestros por colisión en vehículos livianos de la empresa en la ciudad de Guayaquil el periodo 2021-2022

Fuente: Resultados de la investigación.

4.2. Estadística Descriptiva

En base a la información estudiada de los reclamos atendidos en la ciudad de Guayaquil por daños directos por choque, se detallan a continuación los datos estadísticos obtenidos del presente trabajo:

En la Tabla 2 y 3 revisamos los resultados obtenidos con relación a los cuartiles de la muestra, el rango, y los pagos mínimo y máximo registrados en la base histórica de siniestros atendidos en la ciudad de Guayaquil. A su vez también se menciona otras medidas estadísticas como la media, la mediana, la desviación estándar, la curtosis, el sesgo y la cantidad de datos.

La tabla 2 está conformada por los datos estadísticos del diagrama de cajas, que nos ayudan a identificar dónde existe acumulación de datos y cuál es su comportamiento típico dentro de los rangos observados, se aprecia que solamente el 25% son menores a 676,29, que el 50% de los datos están por debajo de 1.466,11. Por otro lado, solo el 25% está por encima de los 3.242,36, lo cual nos permite inferir que la mayor cantidad de los siniestros se encuentran en un rango de cuatro cifras bajas. Adicional, es importante mencionar la presencia de datos atípicos o extremos, donde la pérdida asumida sobrepasa los 100,000.00.

Distribución	Datos
Min (0%)	13,56
Q1 (25%)	676,29
Q2 (50%)	1466,11
Q3 (75%)	3242,36
MAX (100%)	102459,23
RANGO	102445,67
Q3-Q1 (RANGO CUARTÍLICO)	2566,07

Tabla 2.– Distribución cuartiles de severidad 2021-2022

Fuente: Resultados de la investigación

La tabla 3 presenta mediciones de tendencia central en los datos de estudio (la medida y la mediana), mientras que la desviación estándar, sesgo y la curtosis representan los diferentes momentos de la distribución, siendo la desviación un

indicador de dispersión en los datos de la muestra, el valor de un sesgo positivo nos sugiere que el comportamiento de los datos estará sesgado a la derecha, y por último la curtosis obtenida que posee un valor relativamente alto, evidencia que las colas son más largas que las de una distribución normal.

Estadística descriptiva	Datos
Media	3110,12
Mediana	1466,11
Cantidad de Datos	2669
Desviación está.	5303.61
Curtosis	63,93
Sesgo	5,87

Tabla 3 – Estadística descriptiva de Severidad 2021-2022

Fuente: Resultados de la investigación

4.3 Selección de distribución - Severidad

La distribución de Severidad se encuentra modelada a partir de los datos de pérdida monetaria registrada por la compañía en la ciudad de Guayaquil a causa los siniestros atendidos por choque vehículos durante el periodo 2021-2022.

Se observa que la distribución de severidad extraída a partir de la variable de los montos pagados por cada reclamo atendido en 2021-2022, para encontrar la distribución que más se ajuste al modelo, se utilizó el criterio de información Bayesiana y el Akaike, se modela a partir distintas distribuciones paramétricas utilizados en estudios similares tales como (Franco & Murillo, 2008), se utilizó las distribuciones de Lognormal, Inversa de Gauss y Gamma, Weibull, las cuales

funcionan en este tipo de escenarios ya que: permiten modelar casos extremos, presentar distribuciones sesgadas o de cola larga, y resultar en valores decimales.

La distribución más apegada a la real fue la distribución Lognormal, ya que los valores de BIC (47612,72) y AIC (47600,95) fueron los menores en comparación al resto de distribuciones.

La Figura 4 nos permite observar la distribución de la pérdida de los montos pagados por siniestros de choque en Guayaquil, lo que nos permite evidenciar que posee un sesgo a la derecha y que engloba casos de cuantías mínimas y de casos de una cuantía significativamente alta.

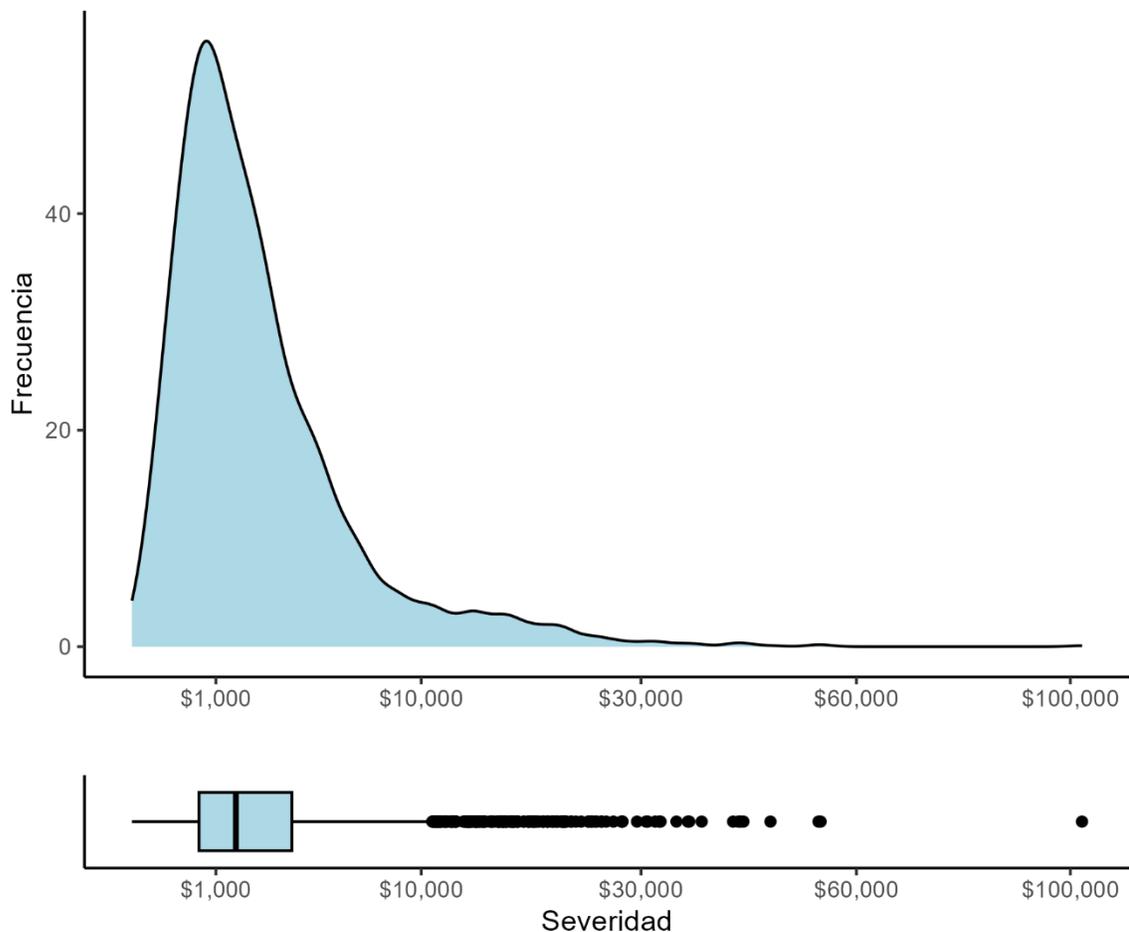


Figura 4. Distribución de severidad 2021-2022

Fuente: Resultados de la investigación.

En la tabla 4, se detallan las distribuciones candidatas para seleccionar la distribución adecuada para el modelo en asunto. Se utilizó los métodos de ajuste sugeridos por la literatura encontrada, siendo estos el criterio Bayesiano, Akaike y K-S. A Continuación, los resultados de las potenciales distribuciones de severidad en los criterios de información BIC y AIC, luego, se procedió a escoger la distribución cuyo valor sea el menor, es decir la Lognormal. En base a la distribución escogida mediante el criterio de información bayesiano, podemos afirmar que el comportamiento de la severidad de los eventos no tiene un comportamiento de sesgo elevado (High Tail) como el descrito por las distribuciones Weibull y Gumbell.

Distribución	GL	BIC	AIC
Lognormal	2	47612,72	47600,95
Weibull	2	48039,25	48027,47
Gumbel	2	50111,43	50099,65
Inversa de Gauss	2	48020,59	48008.81
Gamma	2	48182,25	48170.47

Nota: GL significa grados de libertad

Tabla 4.- Tabla de aplicación del criterio Bayesiano y Akaike en las distribuciones de severidad Lognormal, Weibull, Gumbel, Inversa de Gauss y Gamma

Fuente: Resultados de la investigación.

En la Tabla 5 observamos que el P Value de la prueba K-S es mayor al 5%, por lo tanto, no se rechaza la hipótesis nula, a favor de que la distribución Lognormal describe correctamente el comportamiento de la distribución de severidad de perdidas por choque, mientras que, en las demás distribuciones, la hipótesis nula se descarta

KS	GL	P VALUE	ESTADISTICO
Lognormal	2625	0.188	0.040
Weibull	2625	0.001	0.069
Gumbel	2625	<0.001	0,164
Inversa de Gauss	2625	0.007	0.100
Gamma	2625	0.001	0,12

Tabla 5.- Tabla de aplicación la prueba de Bondad de Ajuste K-S en las distribuciones de frecuencia, Lognormal, Weibull, Gumbel, Inversa de Gauss y Gamma

Fuente: Resultados de la investigación.

En la Figura 5 se realiza comparación por medio de método del Q-Q plot, donde se presentan cuantiles muestrales y teóricos de severidad del q-q plots, cuya función es evidenciar las semejanzas entre la distribución teórica y real de la severidad de los montos pagados por daño directo por choque. Podemos observar que todos los puntos que representan la distribución real de los casos se encuentran dentro del intervalo de confianza de la distribución teórica.

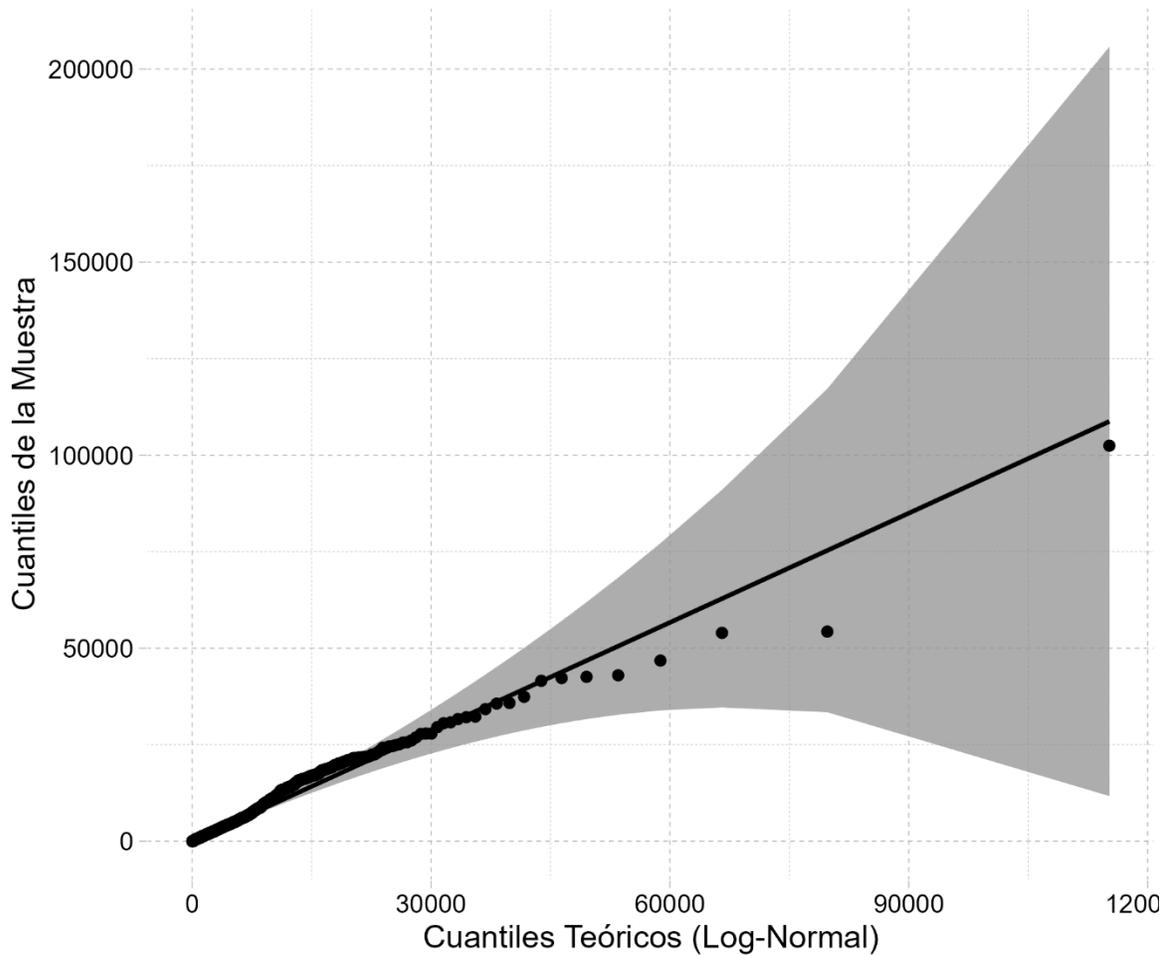


Figura 5.- Ajuste de Severidad con relación a la distribución Lognormal

Fuente: Resultados de la investigación.

A su vez en la Figura 6 se realizó la prueba del Q-Q plot con las demás distribuciones, lo que nos permite visualizar que no se ajustan al modelo, donde se puede apreciar que las distribuciones de Weibull, Gamma y Gumbel tienen un bajo nivel de ajuste con respecto a las diferencias entre los cuantiles teóricos y reales. Además, podemos observar que la distribución Inversa de Gauss es el modelo subóptimo con mejor desempeño.

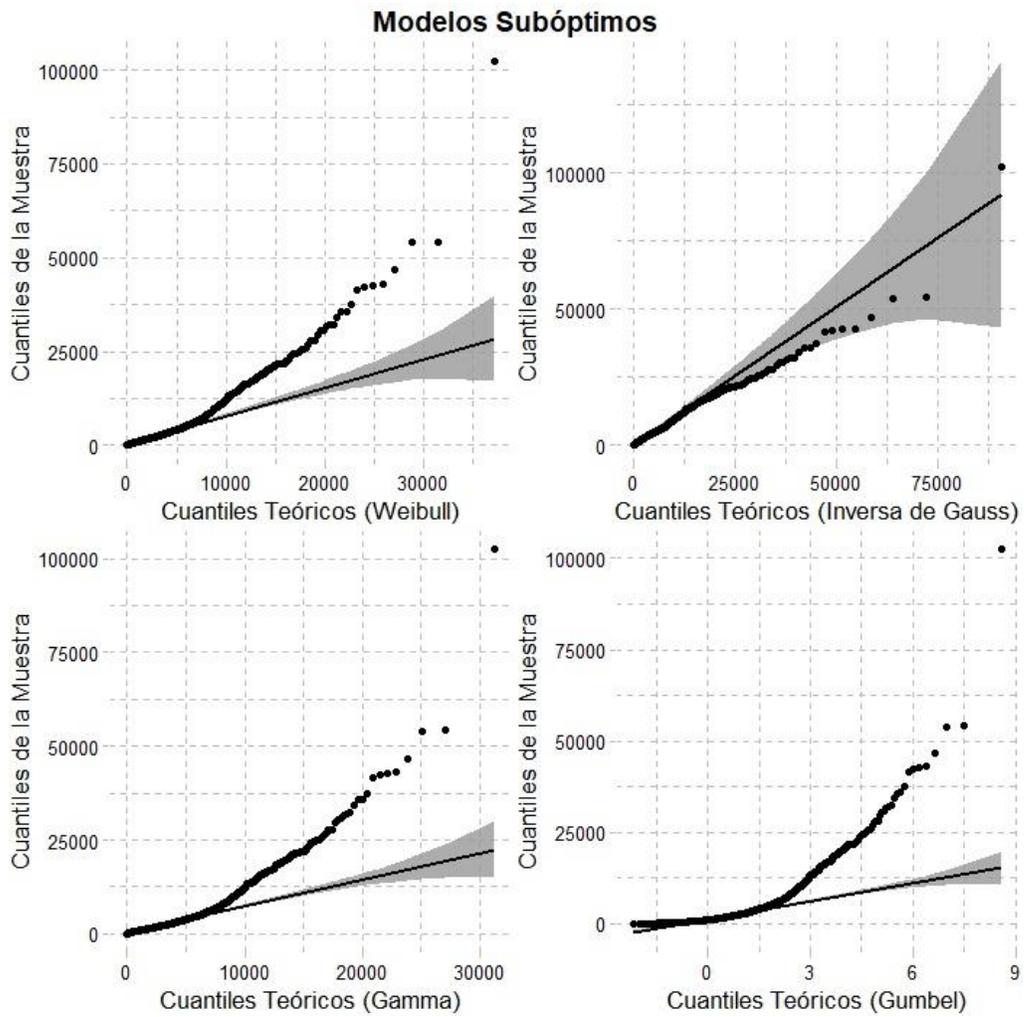


Figura 6.- Ajuste de Severidad con relación a las distribuciones de Weibull, Inversa de Gauss, Gamma y Gumbel

Fuente: Resultados de la investigación.

4.4. Distribución de Frecuencia

La distribución de Severidad se encuentra modelada a partir de la cantidad de siniestros reportados según los datos la compañía NEW LIFE ECUADOR SA en la

ciudad de Guayaquil a causa los siniestros atendidos por choque vehículos durante el periodo 2021-2022.

En la Figura 7 observamos la distribución de frecuencia de los eventos se necesitará como variable la fecha de los eventos reportados registrados en el reporte mensual de siniestros, se realizó las pruebas con la distribución Binomial Negativa y de Poisson ambas siendo variables discretas que por lo tanto nos permiten obtener números enteros, la distribución que finalmente nos brindó mayores bondades de ajuste (un menor BIC y AIC) fue la binomial negativa, Y así mismo, su comportamiento se asemeja mucho más al de la distribución real (Medina, Restrepo, & Bedoya, 2015). De la Figura 7 podemos observar que la distribución de binomial negativa al tener dos parámetros de ajuste a diferencia de la distribución Poisson tiende a ajustarse mejor en términos de mayor varianza, lo cual le permite tener un mejor ajuste.

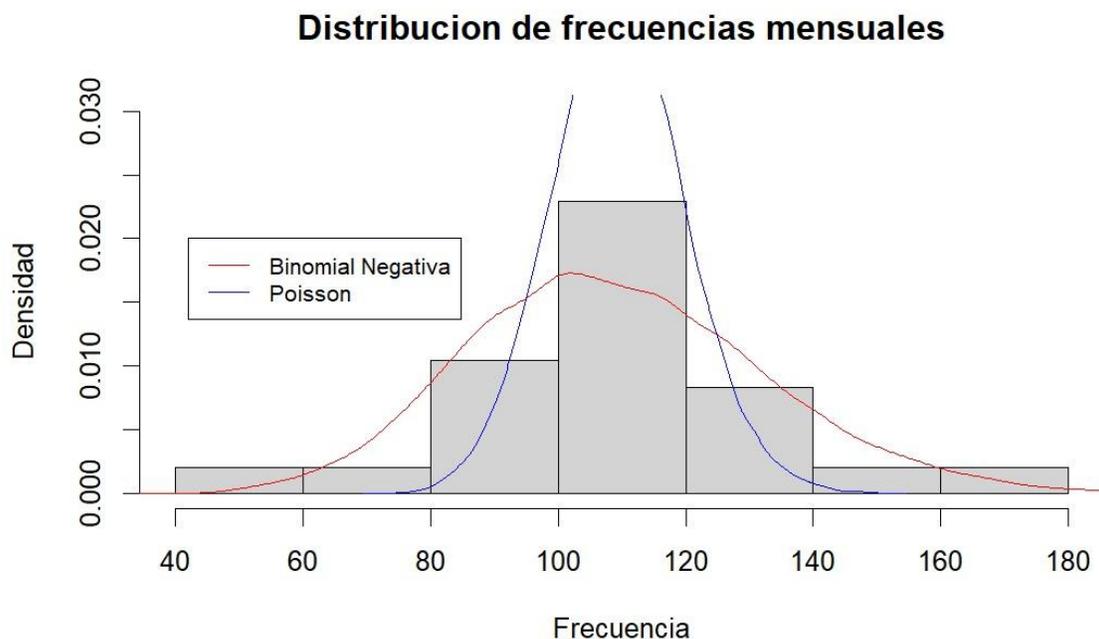


Figura 7.- Distribución de Frecuencia de Reportes atendidos en el periodo 2021-2022

Fuente: Resultados de la investigación.

En la tabla 6 se pueden observar los resultados de las potenciales distribuciones de Frecuencia en los criterios de información BIC y AIC. Luego de escoger la distribución cuyo valor sea el menor, se logró encontrar que la distribución que más se asemeja al comportamiento de la frecuencia de pérdidas por choque era la distribución Binomial Negativa.

Distribución	GL	BIC	AIC
Poisson	1	249,16	247,99
Binomial Negativo	2	218,19	215,84

Tabla 6.- Tabla de aplicación del criterio Bayesiano y Akaike en las distribuciones de frecuencia

Fuente: Resultados de la investigación.

En la Tabla 7 observamos que el P Value de la prueba K-S no supera el umbral del 5%, por lo tanto, no se rechaza la hipótesis nula, a favor de que la distribución Binomial Negativa describe correctamente el comportamiento de la distribución de frecuencia de reportes de siniestros por choque. Por otro lado, se evidencia que la distribución de Poisson también cumpliría dicha Hipótesis, sin embargo, el estadístico KS de la distribución binomial negativa evidencia un mejor desempeño.

KS	GL	P VALUE	ESTADISTICO
Binomial Negativa	24	0.99	0.092
Poisson	24	0.32	0.197

Tabla 7.- Tabla de aplicación la prueba de Bondad de Ajuste K-S en las distribuciones de frecuencia.

Fuente: Resultados de la investigación.

En la Figura 8 utilizamos la gráfica del q-q plots, cuya función es evidenciar las semejanzas entre la distribución teórica y real de la frecuencia de casos por daño directo por choque, dentro de líneas de referencia que actúan como rangos de confianza.

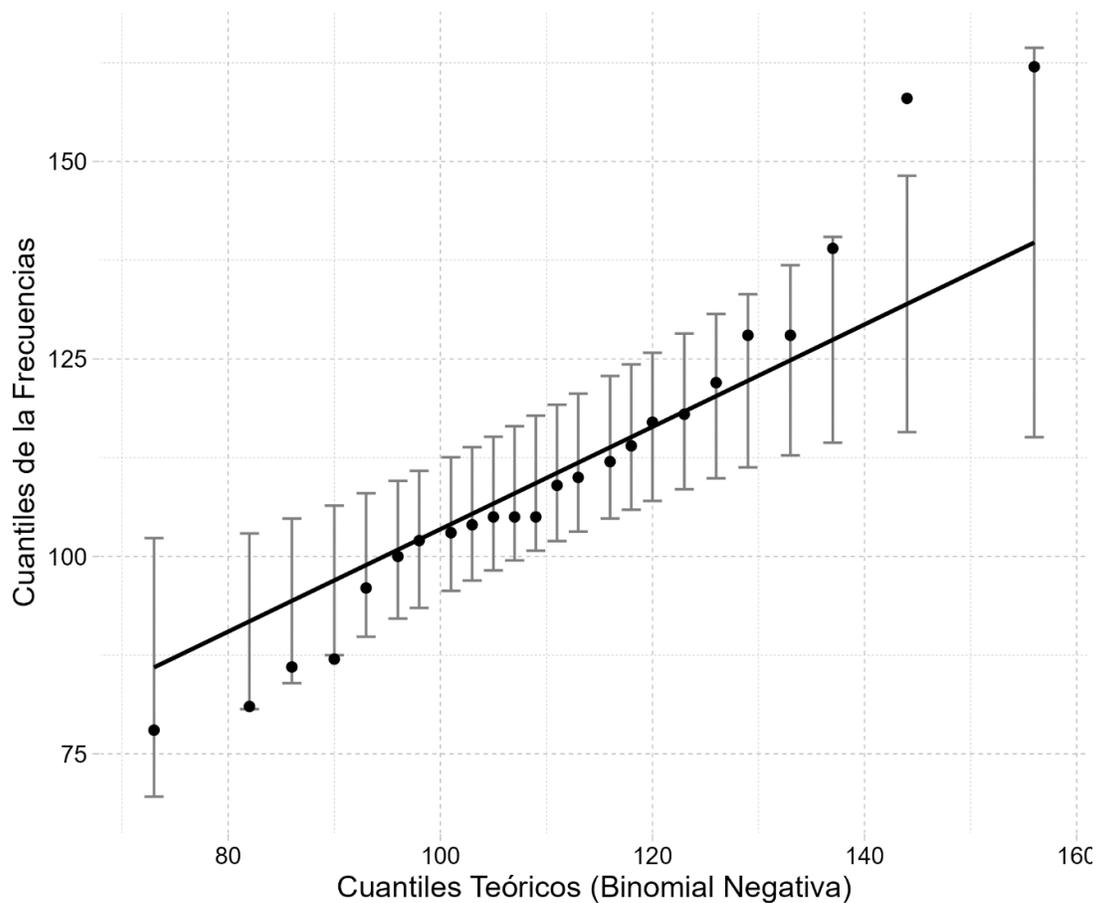


Figura 8.- Ajuste de Frecuencia con relación a la distribución Binomial Negativa

Fuente: Resultados de la investigación.

En la Figura 9 se realizó la prueba del Q-Q plot para la distribución de Poisson, lo que nos permite visualizar que, si bien la mayoría de los puntos se encuentran dentro de las líneas de referencia, estos se presentan mucho mas dispersos, sobre todo si visualizamos sus extremos de mínimo y máximo, por lo tanto, no se considera ideal el uso de esta distribución para el modelo.

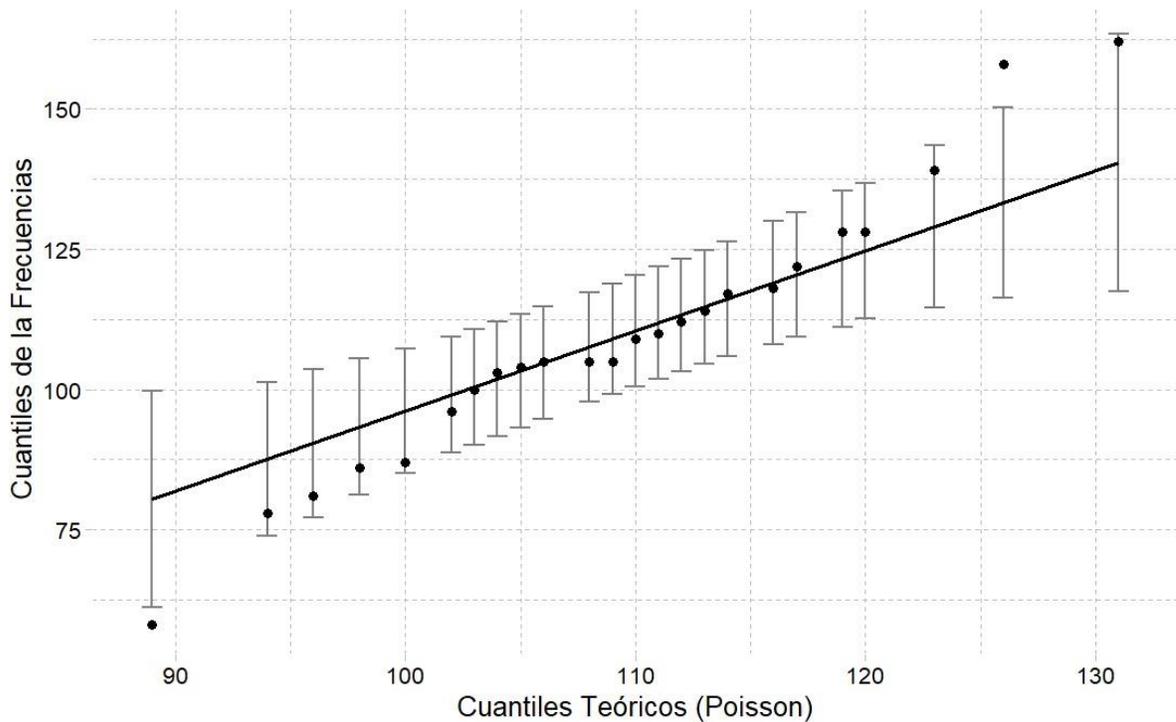


Figura 9.- Ajuste de Frecuencia con relación a la distribución Poisson

Fuente: Resultados de la investigación.

4.5. Estimación de Valor operativo de riesgo (Monte Carlo)

En este apartado se presentará los resultados finales de la investigación una vez realizados los ajustes que permitieron escoger las distribuciones de frecuencia y severidad que más se asemejan a la naturaleza de los datos de estudio.

Dentro de la Figura 9 se remarcan 3 momentos de la distribución de la pérdida, la primera se asocia a la pérdida esperada estimada resultante del modelamiento de los datos, representada por la línea vertical azul, la segunda y tercera indican los valores de la pérdida no esperada en el cuartil 95, diferenciada por el color amarillo que muestran la posibilidad de casos atípicos, y la del cuartil 99, representada por el color rojo, que a su vez representa el valor en riesgo en el cual se deba asumir la pérdida por casos extremos; diversos autores recomiendan el uso de ambos percentiles para un mejor entendimiento del riesgo.

La pérdida esperada resultante de la simulación Monte Carlo por pago de siniestros relacionados a daños directos por choque de vehículos en la ciudad de Guayaquil es de \$ 347.387,00, esto quiere decir que el valor que espera asumir la sucursal dentro de un mes es este valor.

Por otro lado, para el cálculo de Valor en Riesgo, se obtuvo un VaR bajo un nivel de confiabilidad del 95%, la pérdida esperada estará debajo de los \$ 503.987,37, y con un umbral esta vez del 99% de confiabilidad, se posiciona por debajo de los \$ 592.614,1. Siendo estos los indicadores del valor en riesgo al percentil 95 y 99 respectivamente. Se acompañan los presentes resultados de la gráfica con una tabla de datos de los distintos percentiles que nos permiten identificar el posicionamiento de los datos en la distribución de la pérdida. Por otro lado, es importante mencionar que todos los valores que se encuentren posicionados por encima del percentil 99, toman en nombre de escenarios catastróficos.

Percentil	Valor
25%	284792.89
50%	339260.34
75%	400418.87
95% (VaR)	503987.37
99% (VaR)	592614.10
PE	347387.00

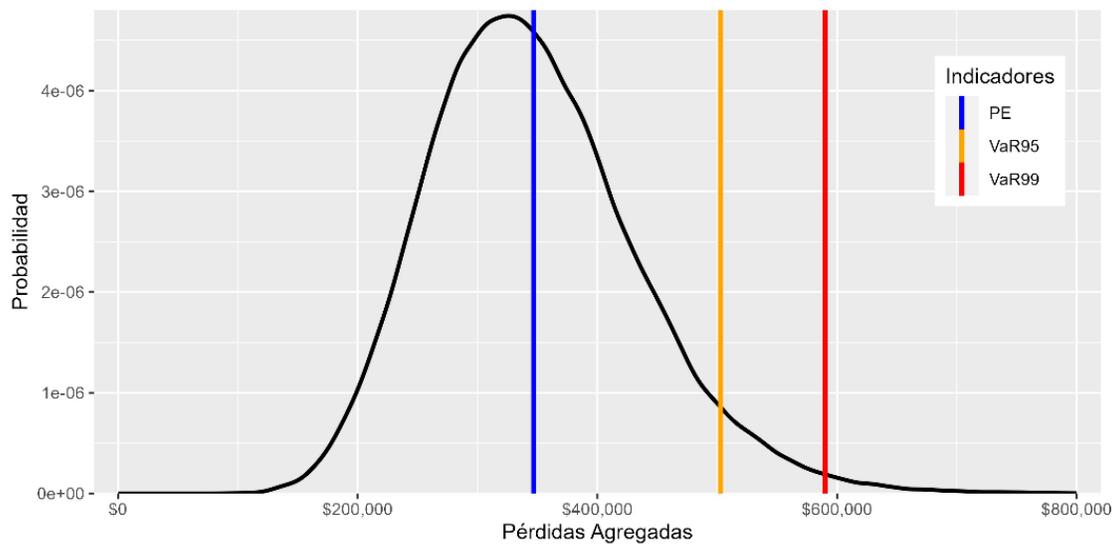


Figura 9.- Distribución de Pérdida Agregada de Siniestros de vehículos livianos de la Empresa NEW LIFE ECUADOR SA. en la ciudad de Guayaquil

Fuente: Resultados de la investigación.

Como podemos observar en la figura 10, la relación de la Pérdida Esperada de Siniestros con la cantidad de simulaciones utilizadas en el método Montecarlo refleja

que conforme aumenta la cantidad de simulaciones, nos aproximamos más al valor de la pérdida esperada obtenido, demostrando así estabilidad alrededor del valor reportado.

Este patrón evidencia que la cantidad escogida de 100000 simulaciones para el método Monte Carlo es apropiada para la naturaleza de los datos y para la obtención de los objetivos planteados en este estudio.



Figura 10.- Convergencia de la Pérdida Esperada mediante simulación Monte Carlo

Fuente: Resultados de la investigación.

Capítulo 5. Estrategias

Dentro del estudio realizado, logramos obtener la aproximación estadística de los montos de pérdida probables por pago de siniestros de vehículos dentro de la Ciudad de Guayaquil, sin embargo, se recomienda expandir este modelo probabilístico a las demás ciudades principales del Ecuador, tales como Quito, Cuenca, Ambato, etc. Este análisis comparativo podría lograr identificar las variables de siniestralidad y comportamiento de los individuos en los distintos lugares del país, y en base a ello, proponer estrategias diferenciadas de suscripción y evaluación de futuros asegurados, esto puede traducirse a nuevos métodos para llevar a cabo inspecciones de riesgo, que permitan conocer mejor a los nuevos negocios y precautelar que no se aseguren individuos con alta probabilidad de choque.

Como otra estrategia, también se sugiere establecer con socios estratégicos un mayor control en los precios de los repuestos adquiridos, conseguir precios referenciales con talleres concertados a cambio de volumen de ingreso de vehículos en su taller, este tipo de alianzas puede disminuir el valor de la factura promedio que la aseguradora está asumiendo por evento, del mismo modo lograr convenios con socios importadores multi marca que permitan establecer precios referenciales, especialmente en los casos donde el vehículo siniestrado no sea comercial y el costo de ciertos repuestos aumente por la falta de oferta.

Además, se debe mencionar que a partir del modelo predictivo realizado, se logró concluir que la mayor cantidad de eventos atendidos se encuentran concentrados en un rango de precio de \$ 1.500 a \$ 3.000 dólares, por lo tanto, se pueden buscar procesos diferenciados en la atención de siniestros de esta cuantía, implementando modalidades que logren optimizar recursos y tiempo, reducir tiempos de aprobación y mitigar los retrasos por operatividad, de esa forma, las personas que conforman el área de siniestros podrán ser más efectivos a la hora de identificar señales de fraude y sobreprecio, disminuyendo así los montos de pérdida y la exposición al riesgo vigente.

Capítulo 6. Conclusiones

Los resultados obtenidos en el presente estudio mediante la simulación Montecarlo basada en el enfoque LDA (Loss Distribución Approach) a partir de información de pagos por siniestros de vehículos en la cobertura de daño directo por choque permiten concluir que el valor de la pérdida esperada para la compañía de seguros NEW LIFE ECUADOR SA es aproximadamente un total de \$ 347.387,00.

Con respecto El VaR o máxima pérdida probable, se obtuvo que al 99% de nivel de confianza es de \$ 592.614,10. este valor debe ser considerado para el cálculo de reservas técnicas y así poder conocer el valor monetario a utilizar en el caso de ocurrir eventos desfavorables. Obteniendo así una Medida clara de la pérdida potencial para la aseguradora a la hora de asumir un riesgo en un escenario de selección adversa. Siendo esta la contribución más importante del estudio. A su vez, en el presente trabajo también se considero el VaR al 75% 95% como indicadores, siendo \$ 284.792,89 \$ 503.987,37 respectivamente.

Desde el punto de vista metodológico se logró una aproximación estadística de la exposición al riesgo presente en el giro de negocio tanto para la frecuencia como para la severidad de los eventos, este estudio se a realizo con la simulación Montecarlo acompañado de tres enfoques distintos de ajuste, lo cual le otorga robustez y veracidad al modelo.

Por otro lado, a partir del análisis exploratorio de los datos, se obtuvo una aproximación estadística que nos permite conocer la distribución real de las perdidas asumidas, revelando una concentración de datos cuyo monto oscila entre los 1500 y 3000 dólares. lo cual permite a la compañía diseñar estrategias de gestión de reclamos enfocados en ciertos rangos de cuantía, y así poder también elaborar procedimientos diferenciados que mejoren los tiempos de respuesta y atención en el área.

Bibliografía

- Almeida, A., Loy, A., & Hofmann, H. (2018). *Contributed Research Articles*, 48. doi: ISSN 2073-4859
- Aslam, M. (2020). Introducing Kolmogorov–Smirnov Tests under Uncertainty: An. *American Chemical Society*, 5, 1, 914–917. doi:<https://doi.org/10.1021/acsomega.9b03940>
- Azaare, J., Wu, Z., Zhu, Y., Armah, G., Engmann, G. M., Kwadwo, S. M., . . . Ampaw, E. M. (2022). Measuring the Adequacy of Loss Distribution for the Ghanaian Auto Insurance Risk Exposure through Maximum Likelihood Estimation. *Open Journal of Business and Management*.
- Bank, J. (1999). *Introduction to Simulation*. Atlanta: AutoSimulations, Inc.
- Barbu, A., & Zhu, S.-C. (2020). *Monte Carlo Methods* (Springer Nature Singapore ed.). Tallahassee, Florida: Springer. doi:<https://doi.org/10.1007/978-981-13-2971-5>
- Bauldry, S. (2015). Structural Equation Modeling. *International Encyclopedia of the Social & Behavioral Sciences*, volume 22, pp. 15215–15222, 615–620. doi: doi:10.1016/b978-0-08-097086-8.44055-9
- Calero, M. (26 de Octubre de 2022). *Actuaría*. Obtenido de Evolución de los principales actores en el mercado de seguros: <https://actuaría.com.ec/es/evolucion-de-los-principales-actores-en-el-mercado-de-seguros/>
- Cevallos, D., & Téran, L. (Julio de 2018). *Implementación del Seguro de Vehículos PPU*. Obtenido de <https://www.dspace.espol.edu.ec/bitstream/123456789/44248/1/D-P13720.pdf>

- Cummins, D., Doherty, N., & Lo, A. (2002). Can insurers pay for the “big one”? Measuring the capacity of the insurance market to respond to catastrophic losses. *Journal of Banking & Finance*, 26, 557–583. doi:[https://doi.org/10.1016/S0378-4266\(01\)00234-5](https://doi.org/10.1016/S0378-4266(01)00234-5)
- De la Fuente, L. (2012). Modelos de Pérdida Esperada. En *Medición Integral del Riesgo Crediticio* (págs. 44-46). Balderas: Limusa.
- Elorriaga, E. R. (2010). *Las nuevas medidas de Basilea III en materia de capital*. Obtenido de Banco de España: <https://www.bde.es/f/webbde/Secciones/Publicaciones/InformesBoletinesRevistas/RevistaEstabilidadFinanciera/10/Nov/Fic/ref0119.pdf>
- Franco, L., & Murillo, J. (2008). LDA Metodología Actuarial Aplicada al Riesgo Operacional. *Revista Ingenierías Universidad de Medellín*, 7, pp. 143-156. doi:1692-3324
- Giraldo, C. A. (2002). Condiciones Generales, Cláusulas Abusivas y el principio de buena fé en el contrato de seguro. *Red de Revistas Científicas de América Latina, el Caribe, España y Portugal*, 233-249. doi: 0041-9060
- Hanafy, M., & Ming, R. (20 de Febrero de 2021). *Learning Approaches for Auto Insurance Big Data*. Obtenido de <https://doi.org/10.3390/risks9020042>
- Hanusz, Z., & Tarasińska, J. (2015). Normalization of the Kolmogorov–Smirnov and Shapiro–Wilk tests of normality. *Biometrical Letters*, Vol. 52 (2015), No. 2, 85 - 93. doi:10.1515/bile-2015-0008
- Heba, M., Cordeiro, G., Haitham, Y., & Mohamed, I. (3 de Agosto de 2021). *A size-of-loss model for the negatively skewed insurance claims data: applications, risk analysis using different methods and statistical forecasting*. Obtenido de <https://www.tandfonline.com/>: <https://doi.org/10.1080/02664763.2022.2125936>

INEC. (Marzo de 2022). *Estadísticas de Transporte*. Obtenido de https://www.ecuadorencifras.gob.ec/documentos/web-inec/Estadisticas_Economicas/Estadistica%20de%20Transporte/ESTRA_2021/2022_ESTRA_SINIESTROS.pdf

Journal of the American Statistical Association. (s.f.).

Klugman, S., Panjer, H., & Willmot, G. (2019). *Loss Models: From Data to Decisions* (Fifth ed.). (S. o. Actuaries, Ed.) Waterloo, United States: 2019 John Wiley and Sons.

Macho-Stadler, I., & Perez-Castrillo, D. (2001). *An Introduction to the Economics of Information: Incentives and Contracts*. (O. U. Press, Ed.) Boston, USA. doi: 0199243255,9780199243259

Macias, G., Parra, S., & Carvajal, E. (2018). Modelo LDA para medición avanzada de riesgo. *Innovar*, 28, 9-27. doi:<https://doi.org/10.15446/innovar.v28n68.70335>

Magri, A., Farrugia, A., Valletta, F., & Grima, S. (2019 de Junio de 2019). *An Analysis of the Risk Factors Determining Motor Insurance Premium in a Small Island State: The Case of Malta*. Obtenido de https://www.um.edu.mt/library/oar/bitstream/123456789/73114/1/An_analysis_of_the_risk_factors_determining_motor_insurance_premium_in_a_small_island_state.pdf

Mazviona, B. W., & Chiduzza, T. (2013). *THE USE OF STATISTICAL DISTRIBUTIONS TO MODEL CLAIMS IN MOTOR INSURANCE*. Obtenido de <https://ijbel.com/wp-content/uploads/2014/01/KLB3140-THE-USE-OF-STATISTICAL-DISTRIBUTIONS-TO-MODEL-CLAIMS-IN-MOTOR-INSURANCE-Batsirai.pdf>

Medina, S., Restrepo, J., & Bedoya, A. (Agosto de 2015). Pérdidas esperadas y detrimento patrimonial por hurto de vehículos en Colombia. *Cuadernos de Economía*, 36(71), 261-292. Obtenido de PÉRDIDAS ESPERADAS Y

DETRIMENTO PATRIMONIAL POR HURTO DE VEHÍCULOS EN COLOMBIA: <http://www.scielo.org.co/pdf/ceco/v36n71/0121-4772-ceco-36-71-00261.pdf>

Oktaviana, P., & Irhamah. (2021). Kolmogorov-Smirnov Goodness-of-Fit test for identifying distribution of the number of earthquakes and the losses due to earthquakes in Indonesia. *Journal of Physics: Conference Series*, 1821 (2021) 012045. doi:10.1088/1742-6596/1821/1/012045

Parze, E. (1988). Akaike Information Criterion Statistics. *Journal of the American Statistical Association*, 907. doi:10.1080/01621459.1988.10478680

Pintado, M. J. (2019). *Aplicación del método Risicar para la gestión de siniestralidad en el ramo vehicular en el sector asegurador ecuatoriano en el año 2018*. Obtenido de <https://dspace.uazuay.edu.ec/bitstream/datos/9382/1/15020.pdf>

Pitt, M., & Myung, I. J. (1 de Octubre de 2002). *When a good fit can be bad*. doi:10.1016/s1364-6613(02)01964-2

Razal, N., & Bee Wah, Y. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics*, Vol.2 No.1, 21-33. Obtenido de : <https://www.researchgate.net/publication/267205556>

Rendón, M. S. (2014). *Modelación de Pérdida Agregada Aplicada a Siniestros de Cáncer de Mama: Caso de una Empresa Aseguradora*. Obtenido de Modelación de Pérdida Agregada Aplicada a Siniestros de Cáncer de Mama: Caso de una Empresa Aseguradora: <http://hdl.handle.net/20.500.11799/68072>

Rendón, M. S. (Febrero de 2014). *Modelación de Pérdida Agregada Aplicada a siniestros de Cáncer de mama: Caso de una empresa aseguradora* . Obtenido de Modelación de Pérdida Agregada Aplicada a siniestros de Cáncer de mama: Caso de una empresa asegurador: <http://ri.uaemex.mx/handle/20.500.11799/68072>

- Ríos, A. R., & Peña, A. M. (2020). Estadística inferencial. Elección de una prueba estadística no paramétrica en investigación científica. *Horizonte de La Ciencia*, 191–208. doi:<https://doi.org/10.26490/uncp.horizonteciencia.2020.19.597>
- Sampieri, R. H. (2017). *Metodología de la Investigación*. Obtenido de <https://www.uca.ac.cr/wp-content/uploads/2017/10/Investigacion.pdf>
- Septiany, R., Setiawaty, B., & Purnaba, G. P. (2020). The use of monte carlo method to model the aggregate loss distribution. *Al-Jabar: Jurnal Pendidikan Matematika*.
- Shevchenko, P. V. (2012). Calculation of aggregate loss distributions. *Journal of Operational Risk*.
- Shi, P., Feng, X., & Ivantsova, A. (20 de Julio de 2015). Claims, Dependent frequency–severity modeling of insurance. *Insurance: Mathematics and Economics*. Obtenido de Insurance: Mathematics and Economics: <http://dx.doi.org/10.1016/j.insmatheco.2015.07.006>
- Sotomayor, S. (2012). *Estimación de la pérdida esperada para una cartera de microcrédito de basada en calificaciones internas*. Obtenido de <http://bibdigital.epn.edu.ec/handle/15000/4668>
- Stiglitz, J. (2002). La información y el cambio en el paradigma de la ciencia económica. *Revista Asturiana de Economía*, 95-164. doi: ISSN 1134-8291
- Superintendencia de Compañías Valores y Seguros. (Septiembre de 2016). *LEY GENERAL SEGUROS*. Obtenido de Corporación del Seguro de Depósitos, Fondo de Liquidez y Fondo de Seguros Privados: <https://www.cosede.gob.ec/wp-content/uploads/2016/08/LEY-GENERAL-SEGUROS.pdf>

- Torri, G., Radi, D., & Hana, D. (Junio de 2022). Catastrophic and systemic risk in the non-life insurance sector: A micro-structural contagion approach. *Finance Research Letters*, 47(b). doi:<https://doi.org/10.1016/j.frl.2022.102718>
- Transportation, C. F. (1999). *Automobile Insurance: Road Safety, New Drivers, Risks, Insurance Fraud and Regulation*. New York: Springer Science+Business Media, LLC. doi:10.1007/978-1-4615-4058-8
- Wang, N., Qian, L., & Liu, N. Z. (2 de September de 2019). *Modelling the aggregate loss for insurance claims with dependence*. Obtenido de Modelling the aggregate loss for insurance claims with dependence: <https://doi.org/10.1080/03610926.2019.1659368>
- Yousof, H., Tashkandy, Y., Emam, W., Ali, M., & Ibrahim, M. (7 de January de 2023). A New Reciprocal Weibull Extension for Modeling Extreme Values with Risk Analysis under Insurance Data. *New Advances and Applications of Extreme Value Theory*. doi:<https://doi.org/10.3390/math11040966>

Anexos

Códigos utilizados en R

```
title: "AED Siniestos"
```

```
author: "Maria Paula Bazurto"
```

```
date: "`r Sys.Date()`"
```

```
output: html_document
```

```
library(openxlsx) # Para tratar excels
```

```
library(dplyr) # deployment
```

```
data = read.xlsx('BASE TESIS.xlsx',detectDates = T)
```

```
data <- janitor::clean_names(data)
```

```
head(data)
```

```
data <- data %>% mutate(tipo_vehiculo = if_else(grepl(x =  
tipo_vehiculo,'LIV'),'LIVIANO',tipo_vehiculo),
```

```
                  tipo_vehiculo = if_else(grepl(x =  
tipo_vehiculo,'PES'),'PESADO',tipo_vehiculo))
```

```
data %>% filter(even_code_name %in% c('DANO DIRECTO POR CHOQUES',
```

```
                                  'PERDIDA PARCIAL POR CHOQUE',
```

```

                                'PERDIDA TOTAL POR CHOQUE')) %>%
filter(sucursal_nombre == 'GUAYAQUIL' &
      tipo_vehiculo == 'LIVIANO') -> datos2
dim(datos2)

datos2 %>% mutate(exp_final = paid_loss_local,
                 even_date = as.Date(even_date)) %>%
group_by(claim) %>%
summarise(
  fecha = min(even_date),
  n_pagos_cobros = n(),
  severidad = sum(exp_final)) -> data3

dim(data3)
data3 %>% filter(fecha >= as.Date('2021-01-01') & fecha < as.Date('2023-01-01') &
                 severidad > 0) -> data4

dim(data4)
library(e1071)
summary(data4$severidad)
print('desviacion standar')
sd(data4$severidad)
print('sesgo')

```

```
skewness(data4$severidad)
```

```
print('curtosis')
```

```
kurtosis(data4$severidad)
```

```
print('cantidad de datos')
```

```
nrow(data4)
```

```
library(e1071)
```

```
data_mensual <- data4 %>%
```

```
  mutate(periodo = as.Date(lubridate::floor_date(fecha,"month"))) %>%
```

```
  group_by(periodo) %>% summarise(n = n(), sev = sum(severidad))
```

```
summary(data_mensual$n)
```

```
print('desviacion standar')
```

```
sd(data_mensual$n)
```

```
print('sesgo')
```

```
skewness(data_mensual$n)
```

```
print('curtosis')
```

```
kurtosis(data_mensual$n)
```

```
print('cantidad de datos')
```

```
nrow(data_mensual)
```

```
severidad <- data4$severidad
```

```
library(univariateML)
```

```
BIC(
```

```
  mllnorm(severidad),
```

```
  mlweibull(severidad),
```

```
  mlgumbel(severidad),
```

```
  mlinvgauss(severidad),
```

```
  univariateML::mlgamma(severidad)
```

```
AIC(
```

```
  mllnorm(severidad),
```

```
  mlweibull(severidad),
```

```
  mlgumbel(severidad),
```

```
  mlinvgauss(severidad),
```

```
  univariateML::mlgamma(severidad)
```

```
)
```

```

plot(mllnorm(severidad), col = "red")

lines(density(severidad,bw = 90))

set.seed(12)

qqplot(y = severidad,x = rml(10000,mllnorm(severidad)))

abline(a=0, b=1)

require(qqplotr)

qqplotsev <- ggplot(data = data4, mapping = aes(sample = severidad)) +
  stat_qq_band(distribution = 'lnorm') +
  stat_qq_line(distribution = 'lnorm') +
  stat_qq_point(distribution = 'lnorm') +
  labs(x = "Cuantiles Teóricos (Log-Normal)",
       y = "Cuantiles de la Muestra") +
  theme_pander()

ks.test(severidad,x = rml(10000,mllnorm(severidad)))

ggsave('graficas/qqseveridad.png',plot = qqplotsev,
       width = 6, height = 5,units = 'in')

```

```

library(lubridate)

data_mensual <- data4 %>%
  mutate(periodo = as.Date(lubridate::floor_date(fecha,"month"))) %>%
  group_by(periodo) %>% summarise(n = n(), perdida = sum(severidad))

frecuencia <- data_mensual$n

hist(frecuencia, main = "Distribucion de frecuencias mensuales",
      probability = TRUE,ylab = 'Densidad',ylim = c(0,0.030), xlab = 'Frecuencia')
lines(density(rnbinom(10000,size = nbinomial[[1]][[1]],
                    mu = nbinomial[[1]][[2]])), col = 'red')
lines(density(rpois(10000,mean(data_mensual$n))), col = 'blue')
legend(42, 0.02, legend=c("Binomial Negativa", "Poisson"),
      col=c("red", "blue"), lty=c(1,1), cex=0.8)

library(actuar)
library(fitdistrplus)

nbinomial <- fitdist(frecuencia, "nbinom")
poisson <- fitdist(frecuencia, "pois")

qqcomp(nbinomial)
qqcomp(poisson)

```

```
ks.test(y = frecuencia,x = rnbinom(1000,size = nbinomial[[1]][[1]],  
mu = nbinomial[[1]][[2]])
```

```
qqplotfreq <- ggplot(data = data_mensual, mapping = aes(sample = n)) +  
  stat_qq_band(distribution = 'nbinom',conf = 0.99) +  
  stat_qq_line(distribution = 'nbinom') +  
  stat_qq_point(distribution = 'nbinom') +  
  labs(x = "Cuantiles Teóricos (Binomial Negativa)",  
y = "Cuantiles de la Frecuencias") +  
  theme_pander()
```

```
ggsave('graficas/qqfrecuencias.png',plot = qqplotfreq,  
width = 6, height = 5,units = 'in')
```

```
data.frame(dist = c('pois','binom_neg'),  
  AIC = c(AIC(fitdistr(frecuencia,'poisson'),  
    AIC(fitdistr(frecuencia,'negative binomial'))),  
  BIC = c(BIC(fitdistr(frecuencia,'poisson'),  
    BIC(fitdistr(frecuencia,'negative binomial'))))
```

```
modelo_sev <- mllnorm(severidad)
```

```
nbin <- fitdist(frecuencia, "nbinom")
```

```

folds <- 100

perdidas_finales <- c()

for (i in 1:folds){

  n_simul <- 1000

  ceros = function(x) {

    multiplier <- c(rep(1,x[1]), rep(0,length(x) -1-x[1]))

    x <- x[-1]*multiplier

  }

  eventos <- rbinom(1000,size = nbin[[1]][[1]],

                  mu = nbin[[1]][[2]])

  max_eventos <- max(eventos)

  perdidas_individuales <- rml(n_simul * max_eventos, modelo_sev)

  perdida <- matrix(perdidas_individuales, ncol = max_eventos)

  perdida_matriz <- cbind(eventos,perdida)

  perdidas <- t(apply(perdida_matriz, 1, ceros))

  perdidas_agregadas <- apply(perdidas, 1, sum)

  perdidas_finales <- c(perdidas_finales,perdidas_agregadas)

}

perdidas_nbinom <- perdidas_finales

summary(perdidas_nbinom)

```

```

p3 <- ggplot(data.frame(x = perdidat_nbinom), aes(x = x)) +
  geom_density(size = 1) +
  geom_vline(aes(xintercept = mean(perdidat_nbinom),
    color = "PE"), size=1.2) +
  geom_vline(aes(xintercept = quantile(perdidat_nbinom, c(0.99)),
    color = "VaR99"), size=1.2) +
  geom_vline(aes(xintercept = quantile(perdidat_nbinom, c(0.95)),
    color = "VaR95"), size=1.2) +
  scale_x_continuous(limits = c(0,800000), expand = c(0, 20000),
    labels = scales::dollar_format()) +
  scale_y_continuous(limits = c(0, 5e-06 - 2e-07),
    expand = c(0, 0)) +
  xlab('P rdidas Agregadas') + ylab('Probabilidad') +
  scale_color_manual(name = "Indicadores", values = c(PE = "blue",
    VaR99 = "red",
    VaR95 = 'orange')) +
  theme(legend.position= c(0.9,0.75))

ggsave('graficas/lda.png',plot = p3,dpi = 400,
  width = 8, height = 4,units = 'in')

```

```

perc <- c(0.25, 0.5, 0.75, 0.8, 0.85, 0.9, 0.95, 0.98, 0.99, 0.999)
distr <- c(quantile(perdidas_nbinom, probs = perc),
          mean(perdidas_nbinom))
Percentile <- c(perc, "Mean")
tablafinal <- data.frame(Percentile, distr)
tablafinal
distr <- c(quantile(perdidas_nbinom, probs = perc),
          mean(perdidas_nbinom))
Percentile <- c(perc, "Mean")
tablafinal <- data.frame(Percentile, distr)
tablafinal
mean(perdidas_nbinom) # PE
quantile(perdidas_nbinom, c(0.999)) # VaR
quantile(perdidas_nbinom, c(0.999)) - mean(perdidas_nbinom) # PI
prom = c()
for(i in seq(100,99900,100)){
  if(1 < i){
    localdata = perdidas_nbinom[1:(i+100)]
    prom = c(prom, mean(localdata))
  }
}

```

```

simuls = data.frame(simulaciones = seq(100,99900,100),
                    pe = prom)

simuls_plot <- simuls %>% ggplot(aes(x = simulaciones, y = pe)) +
  geom_line() + scale_y_continuous(labels = scales::dollar_format()) +
  geom_hline(yintercept = mean(perdidas_nbinom),color = "red", size = 1.1) +
  ggtitle("Estimacion de la Perdida Esperada Mensual",
          subtitle = "Utilizando 100000 simulaciones") +
  theme_bw() + ylab("Perdida Esperada") + xlab("Nro Simulaciones")

#simuls_plot

ggsave(filename = "graficas/simuls.png",plot = simuls_plot,width = 6,height = 4.5)

require(ordinal)
require(ggpubr)

BIC(
  mllnorm(severidad),
  mlweibull(severidad),
  mlgumbel(severidad),
  mlinvgauss(severidad),
  univariateML::mlgamma(severidad)
)

ggplot(data = data4, mapping = aes(sample = severidad)) +
  stat_qq_band(distribution = 'lnorm') +
  stat_qq_line(distribution = 'lnorm') +

```

```

stat_qq_point(distribution = 'lnorm') +
labs(x = "Cuantiles Teóricos (Log-Normal)",
     y = "Cuantiles de la Muestra") +
theme_pander()
p1 <- ggplot(data = data4, mapping = aes(sample = severidad)) +
  stat_qq_band(distribution = 'weibull') +
  stat_qq_line(distribution = 'weibull') +
  stat_qq_point(distribution = 'weibull') +
  labs(x = "Cuantiles Teóricos (Weibull)",
       y = "Cuantiles de la Muestra") +
  theme_pander()
p2 <- ggplot(data = data4, mapping = aes(sample = severidad)) +
  stat_qq_band(distribution = 'invgauss', dparams = list(mean =
3110.1,shape=840.7)) +
  stat_qq_line(distribution = 'invgauss', dparams = list(mean =
3110.1,shape=840.7)) +
  stat_qq_point(distribution = 'invgauss', dparams = list(mean =
3110.1,shape=840.7)) +
  labs(x = "Cuantiles Teóricos (Inversa de Gauss)",
       y = "Cuantiles de la Muestra") +
  theme_pander()
p3 <- ggplot(data = data4, mapping = aes(sample = severidad)) +

```

```

stat_qq_band(distribution = 'gamma', dparams =
list(shape=0.7939520,rate=0.0002553)) +

stat_qq_line(distribution = 'gamma', dparams =
list(shape=0.7939520,rate=0.0002553)) +

stat_qq_point(distribution = 'gamma', dparams =
list(shape=0.7939520,rate=0.0002553)) +

labs(x = "Cuantiles Teóricos (Gamma)",
      y = "Cuantiles de la Muestra") +
theme_pander()

p4 <- ggplot(data = data4, mapping = aes(sample = severidad)) +
stat_qq_band(distribution = 'gumbel') +
stat_qq_line(distribution = 'gumbel') +
stat_qq_point(distribution = 'gumbel') +
labs(x = "Cuantiles Teóricos (Gumbel)",
      y = "Cuantiles de la Muestra") +
theme_pander()

plot <- ggpubr::ggarrange(p1,p2,p3,p4,ncol = 2,nrow = 2)

png(width = 700, height = 700, res = 100)
ggpubr::annotate_figure(plot, top = text_grob("Modelos Subóptimos",
color = "black", face = "bold", size = 14))

```

```
dev.off()
```

```
options(scipen = 100)
```

```
ks.test(severidad,rml(1000,mlweibull(severidad)))
```

```
ks.test(severidad,rml(1000,mlinvgauss(severidad)))
```

```
ks.test(severidad,rml(1000,mlgumbel(severidad)))
```

```
ks.test(severidad,rml(1000,univariateML::mlgamma(severidad)))
```

```
ks.test(severidad,rml(1000,mllnorm(severidad)))
```

```
library(tidyverse)
```

```
library(ggthemes)
```

```
ggplot(data = data_mensual, mapping = aes(sample = n)) +
```

```
  stat_qq_band(distribution = 'pois',conf = 0.99) +
```

```
  stat_qq_line(distribution = 'pois') +
```

```
  stat_qq_point(distribution = 'pois') +
```

```
  labs(x = "Cuantiles Teóricos (Poisson)",
```

```
       y = "Cuantiles de la Frecuencias") +
```

```
  theme_pander()
```

```
ks.test(data_mensual$n,y = rpois(1000, mean(data_mensual$n)))
```

