

Salazar García, Juan Fernando; Guzmán Aguilar, Diana Sirley; Hoyos Nieto, Daniel Arturo

Article

Modelación de una prima de seguros mediante la aplicación de métodos actuariales, teoría de fallas y Black-Scholes en la salud en Colombia

Revista de Métodos Cuantitativos para la Economía y la Empresa

Provided in Cooperation with:

Universidad Pablo de Olavide, Sevilla

Suggested Citation: Salazar García, Juan Fernando; Guzmán Aguilar, Diana Sirley; Hoyos Nieto, Daniel Arturo (2023) : Modelación de una prima de seguros mediante la aplicación de métodos actuariales, teoría de fallas y Black-Scholes en la salud en Colombia, Revista de Métodos Cuantitativos para la Economía y la Empresa, ISSN 1886-516X, Universidad Pablo de Olavide, Sevilla, Vol. 35, pp. 330-359, <https://doi.org/10.46661/revmetodoscuanteconempresa.5800>

This Version is available at:

<https://hdl.handle.net/10419/315378>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by-sa/4.0/>



Modelación de una prima de seguros mediante la aplicación de métodos actuariales, teoría de fallas y Black-Scholes en la salud en Colombia

SALAZAR GARCÍA, JUAN FERNANDO

Universidad de Medellín (Colombia)

Correo electrónico: juanfersalazarg@hotmail.com

GUZMÁN AGUILAR, DIANA SIRLEY

Universidad de Medellín (Colombia)

Correo electrónico: dsguzman@udemedellin.edu.co

HOYOS NIETO, DANIEL ARTURO

Universidad de Medellín (Colombia)

Correo electrónico: dahoyos@gmail.com

RESUMEN

La prima de la tarificación en un seguro para el sector salud está influenciada por la siniestralidad de sus suscriptores, lo que genera altos niveles de fluctuación e incertidumbre. El objetivo de esta investigación es la aplicación de los modelos actuariales riesgo individual, riesgo colectivo y modelo de credibilidad, junto con la aplicación del modelo tecnológico de tasa de falla y el modelo de opciones financieras de Black-Scholes como herramientas de estimación de la prima de la tarificación para la industria aseguradora y de la salud en Colombia. A partir de las reclamaciones y de los costos totales de los siniestros históricos se aplican los modelos que permita asegurar primas óptimas para una cobertura a las pérdidas agregadas de los siniestros. Al final, se comparan dichos modelos y se aproxima a una definición de un método óptimo. La importancia de la investigación radica en el alto compromiso, responsabilidad e incidencia financiera de gestionar y mitigar el impacto del riesgo actuarial, planteando nuevas metodologías mediante un nivel de estimación óptima en las primas para certificar un correcto funcionamiento a las entidades del sector en temas de costos, sostenibilidad y cumplimiento al servicio en el sector.

Palabras clave: prima de seguros; estudios actuariales; modelo de riesgo individual; modelo de riesgo colectivo; tasa de fracaso; teoría de la credibilidad; valoración de activos financieros; monto total de reclamaciones.

Clasificación JEL: D91, G12, G22, G52, I12.

MSC2010: 62P05, 60E05.

Modelling of an insurance premium through the application of actuarial methods, failure theory and Black-Scholes in the health in Colombia

ABSTRACT

The pricing's premium in an insurance for the health sector is influenced by the claims ratio of its subscribers, which generates high levels of fluctuation and uncertainty. The objective of this research is the application of the actuarial individual risk models, collective risk and credibility model, together with the application of the failure rate technological model and the Black-Scholes financial options model as tools for estimating pricing's premium for the insurance and health industry in Colombia. Based on the claims and the total costs of the historical claims, the models are applied to ensure optimal premiums for coverage of the aggregate losses of the claims. In the end, comparing these models and approaching a definition of an optimal method. The importance of the research settles in the high commitment, responsibility and financial impact of managing and mitigating the impact of actuarial risk, proposing new methodologies through an optimal estimation level in premiums to certify proper functioning of the sector entities in matters of costs, sustainability and service compliance in the sector.

Keywords: insurance premium; actuarial studies; individual risk model. collective risk model; failure rate; credibility theory; valuation of financial assets; aggregate claim amount.

JEL classification: D91, G12, G22, G52, I12.

MSC2010: 62P05, 60E05.



1. Introducción

En la práctica de los seguros, los actuarios mantienen la tarea principal de determinar unas primas adecuadas que permitan cubrir los riesgos que lo asocian. Por una parte, las primas cobradas deben ser suficientes para cubrir prospectivamente los pagos futuros que se derivan de futuras reclamaciones y permita garantizar un funcionamiento estable de la compañía. Por otra parte, las primas no pueden ser superiores a la expectativa del mercado, de tal manera que el producto no pierda competitividad. Por consiguiente, el mayor reto de los riesgos inherentes de los seguros debe ser determinar las primas de la tarificación de tal manera que reflejen fielmente las características, contengan esos riesgos y puedan disiparlos en el contenido del precio. Han existido una variedad de metodologías para cumplir con el propósito del cálculo de prima en los seguros para que pueda cubrir dichos riesgos y cumplir dichos objetivos. Para la fijación de los precios, las técnicas actuariales recurren a metodologías adaptando el estudio de las reclamaciones del asegurado y de las características del riesgo en particular de la cartera. Para ello, introducen metodologías acordes con la información que suministra la historia, y, valiéndose de ello, se utilizan métodos como el riesgo individual, el riesgo colectivo, la técnica de la credibilidad, entre otros modelos actuariales.

El presente documento, fusiona conceptos y metodologías del área de las finanzas con los seguros para enfatizar en lo que concierne al cálculo de la prima de la tarifa. Para ello podemos destacar una de las teorías de las finanzas, conocida como la estrategia al equilibrio financiero, para lo cual, en los seguros medirá mediante ratios e indicadores los incrementos o variaciones entre los ingresos y los costos para el sector asegurador, o, el sector de la salud en pro del aseguramiento en la salud. Evaluar el incremento de las tarifas tiene importancia para controlar el equilibrio y es la razón por la cual se debe de apoyar en unos principios básicos. Es así como la variación de la prima de la tarifa debe de soportarse del principio técnico de *equidad económica* que correlaciona positivamente la existencia entre la tarifa y la cartera. Además, al principio de *suficiencia*, por el que la tarifa cubre razonablemente la tasa de riesgo y los costos propios de la operación. También, ser el producto de la utilización de información estadística que cumpla la exigencia de *homogeneidad*. Y, por último, ser el producto de la utilización de información estadística que cumpla la exigencia de *representatividad*, situaciones descritas en el artículo 26 del Decreto 1570 de 1993, modificado por el artículo 10 del Decreto 1486 de 1994 en Colombia.

La prima del seguro se reconoce como pura, porque en ella aglomera cada uno de los costos asociados al siniestro. La tarifa del seguro, asocia ese precio que le permite a la empresa aseguradora percibir unos ingresos, este a su vez debe contener el costo puro del siniestro. Así mismo, las primas obtenidas deben representar la mejor estimación para el monto de las tarifas de los siniestros. Ya que la probabilidad de ocurrencia de estos siniestros, asocia los tipos de riesgos asumidos a los que está expuesta la población suscrita en una cartera o en un plan de salud. La finalidad encaminada para que la entidad aseguradora pueda mantener la rentabilidad mínima requerida para permanecer en el mercado y satisfacer a los suscritos. Siendo así que, cobrando tarifas suficientes, estos precios netos permitan cubrir las obligaciones de los siniestros, así como las derivadas a la operación, a la administración y a la utilidad para empresas aseguradoras y del sector salud. Ahora, lo anterior muestra el detalle interno de las empresas a la hora de estimar una prima, pero esto debe contener un acercamiento a los acontecimientos externos de una compañía. El concepto de riesgo en el sector Salud en Colombia se introdujo en el artículo 14 de la ley 1122 de 2007 para todas las entidades responsables al aseguramiento en la salud. El *riesgo en salud* la Organización Mundial de la Salud (OMS) lo define como “los factores que incrementan la probabilidad que se presenten resultados adversos para la salud” (2009 p. V).

Una tarifa óptima debe considerar los *objetivos internos* de la empresa entre los cuales la dirección administrativa y financiera debe tener en cuenta el nivel de beneficios, un nivel de crecimiento y un estimado de permanencia a largo plazo para lo que es necesario una solvencia y liquidez. Estos factores nombrados anteriormente determinan y condicionan la tarifa. Además, la empresa debe considerar *objetivos externos*, ya que la empresa opera en un entorno propio del sector, regido en un marco legal y económico, que actúan como restricciones a la hora de operar y planificar su actividad empresarial. Debido a esto, la tarificación la determinan tanto las actuaciones de la empresa como el

entorno político, legal y económico que la rodea; junto con acontecimientos sobrenaturales y morbilidades comunes y atípicas de la sociedad.

La perspectiva dentro del sector salud en la historia muestra notables desafíos. Uno en particular son los retos que tiene que asumir ante los eventos naturales inesperados del mundo. Para ello, las empresas a través de sus mediciones e indicadores registran como pueden verse afectadas e incluso sufrir pérdidas considerables por una serie de factores de riesgos mencionados por (Denuit, Kiriliouk, & Segers, 2015, pág. 163), y determinan la forma como se deben atender las emergencias sanitarias. Existen eventos que son tan pocos probables que exigen una buena estructura en la producción y atención del servicio, además, unos diseños metodológicos bien distribuidos, de tal manera, que puedan incluir estos eventos en sus modelos.

Ahora, en un entorno económico poco favorable y donde las empresas que ofrecen productos y servicios en el mercado de seguros operan bajo condiciones muy competitivas en las instituciones del sector salud (Gulumser, Tonkin, & Johannes, 2002), se tiene la responsabilidad de atraer nuevas suscripciones, no solamente con unos precios atractivos que estén a la medida de la situación económica, además, tarifando con un nivel técnico de suficiencia, para que no esté por debajo de valor mínimo de sostenibilidad para la empresa y para un cubrimiento mínimo de la siniestralidad.

Es así, que el acceso al servicio debe estar en disponibilidad para todo aquel posible suscriptor en la salud independiente de los momentos de estabilidad económica de un país o también en aquellos tiempos donde el riesgo económico es mayor por los inevitables cambios en los ciclos financieros y sociales afectando en todos los sectores demográficos y estratificaciones como factores de riesgo previsible (OCDE, 2016, pág. 46). Por lo que el cálculo óptimo de tarifas permite subsanar una necesidad al compromiso y responsabilidad social en base al aseguramiento de la salud voluntaria.

Si bien, se presentó un acercamiento general a la razón e importancia del cálculo de la prima, adicional a la necesidad del presente artículo que se desarrolla en base a estimar tarifas óptimas y suficientes para el sector, a priori, de tener información confiable de carácter homogéneo, representativo y suficiente para el análisis con el fin de evitar sesgos en los resultados. Rescatando esos beneficios que tiene la atención de los planes voluntarios para el sector salud, de tal manera que evita aglomeraciones en puntos de las instituciones prestadoras de la salud, además de propagación de enfermedades que ameritan un confinamiento generalizado según la gravedad de la morbilidad. Finalmente, en la investigación, son las reclamaciones por la atención en consultas médicas los eventos que se analizarán. Así, la prima del contrato en particular debe tener en cuenta la edad del suscrito expuesto y las distintas categorías de eventos, por lo que el actuario centra su atención en la descripción probabilística de todas las incertidumbres involucradas en los parámetros asociados en el proceso (Migon & Moura, 2005).

Tener una amplia gama de opciones estadísticas para aplicar las metodologías en los eventos en seguros, más precisamente para las primas o tarifas cobradas a los usuarios por conceptos de pólizas, seguros o servicios prepagados en la salud es importante para las empresas que son obligadas a estimar dichos precios. Tener tantas opciones posibles permite no solo al actuario conocer más el negocio por los detalles abarcados en las diferentes metodologías, sino que también permite en el estudio y en los cálculos precisar en los resultados de la tarificación de las empresas de seguros que hace esencial para el cubrimiento de los siniestros de los suscriptores, para el cumplimiento de las obligaciones de la empresa y para el cumplimiento a los socios como tal.

Es por esto que el objetivo del presente artículo se enmarca en la aplicación de cinco modelos estadísticos para la aplicación en el cálculo de la prima en un seguro para el sector salud. Se aplicará el método de riesgo individual, el método de riesgo colectivo, el método de la credibilidad, el método de tasa de fallas de Weibull y modelo de Black-Scholes de opciones Reales estudiados en la prestación en un plan voluntario en salud para predecir las primas de riesgo.

A continuación, el artículo se estructura en primer lugar con la revisión literaria de los modelos. En segundo lugar, se enfoca en cómo modelar la estrategia de primas óptima mediante modelos

tradicionales de tarificación para una compañía de seguros; el apartado explicará los modelos actuariales tradicionales que se aplicarán en el documento. En tercer lugar, cómo se relaciona esta estrategia con modelos de otras ramas aplicables a la tarificación, se presentará dos modelos aplicados en otras ramas ajustándolos a un modelo para la prima de la tarifa. En cuarto lugar, se definen los criterios de aceptación de los modelos. Finalmente, se presentarán los resultados obtenidos y las conclusiones a tener en cuenta.

2. Revisión de la literatura

El *primero de los modelos* es el método de riesgo individual (Garrido & Genest, 2016, p. 120). Es utilizado para el presente artículo como uno de los métodos tradicionales actuariales en la estimación de la tarificación (Minsalud, 2020, pág. 26). Uno de los principales objetivos es la medición de las aproximaciones a la distribución de pérdidas agregadas de siniestros en este método (Roos, 2007, p. 403), donde da el punto de partida en el enfoque de la modelación de la variable reclamaciones o de la variable severidad, para lo que en este método solo una de las dos variables será modelizada mientras que la otra será una constante. En la práctica, inicialmente se centra el análisis en el modelo clásico de riesgo individual, donde la independencia se asume debido a que los contratos son diferentes en el futuro, tanto en cobertura como en exposición al riesgo (Yang, Zhou, & Zhang, 2005, p. 58).

El *segundo de los modelos* es el método de riesgo colectivo (Lee & Shi, 2019, p. 118). Del mismo modo que en el modelo anterior, es común suponer que el número de reclamaciones y sus montos en cada reclamación son independientes. Entonces, en el modelo de riesgo colectivo para diferentes estructuras de dependencia se considera vincular el número de reclamaciones a cada monto de los siniestros para cuantificar la pérdida total de las reclamaciones.

En otros modelos que se estudiaron para el riesgo colectivo, no presentar una independencia en los dos parámetros, presenta una falta de robustez a la hipótesis, por lo que se conjetura la existencia de la dependencia en los parámetros del modelo de riesgo colectivo (Martel, Hernández, & Vázquez, 2012, p. 1430). Además, sus observaciones están en la dependencia (Hua, 2015) donde la cola entre la frecuencia y la severidad están en un contexto de utilización de la atención médica (Lee & Shi, 2019, p. 115).

El *tercer modelo* de la credibilidad es la teoría que pretende enfrentar los problemas de heterogeneidad existente en las carteras, esto es, ser lo más cauto al momento de cobrar lo justo a cada suscriptor, medido así, por el perfil de riesgo asociado que esté presente a nivel individual según sus características independientes que lo identifican (Chen, Cheung, Choi & Yam, 2020, p. 216). Observando el comportamiento de una cartera (plan o portafolio), en la atención de una cartera voluntaria en salud para un suscriptor desde el punto de vista tarifario, a un suscriptor que no haya tenido ningún siniestro en un periodo no debería ser asignado con una prima cero, así como un suscriptor no se le debería penalizar cobrando un valor muy alto en la tarifa por haber sufrido muchos siniestros en un periodo. Esto motiva que, la prima deba ser medida por dos eventos particulares: el primero por la experiencia asociada, en particular, al suscriptor, y, el segundo, por la experiencia de la cartera a la que pertenece (Moreno & Ramos, 2003, p. 30).

Así mismo, la credibilidad se define como “una colección de ideas concernientes al ajuste sistemático de las primas de seguros a medida que se obtiene la experiencia de reclamaciones” (Miller & Hickman, 1975, p. 161). De igual manera se puede interpretar como “la investigación de ciertos principios y métodos para ajustar las primas a medida que la experiencia de las reclamaciones se obtiene” (Norberg, 1979, p. 181). Ambas definiciones dan claridad a la razón de ser de una de las metodologías para tener en cuenta y manifiestan la importancia de la historia en el ajuste de las primas. Por supuesto, se debe tener en cuenta uno de los principios de la actuaría, el cual es, los datos deben cumplir el criterio de suficiencia. Para el modelo de la credibilidad, varios autores han tomado como punto de partida esta metodología. A continuación, veamos algunos.

El modelo clásico de Bühlmann (1969). Se determina un estimador lineal que permite ponderar la experiencia individual con todo el plan (la cartera). Este modelo considera la siniestralidad y no requiere un nivel de detalle de información, es decir, si se cuenta con el monto en unidades monetarias de los siniestros no se requiere la información detallada de un tipo de riesgo específico (grupo etario, por género o por morbilidad). Este modelo es el más sencillo en cuanto a su aplicación y requerimientos de datos. Para este modelo es necesario unos estimadores (Moreno & Ramos, 2003) para obtener el factor de credibilidad “z” que se aplican sobre la experiencia en la siniestralidad.

El modelo Bühlmann - Straub (Wen, Wu & Zhou, 2009, p. 72). Se basa en el modelo anterior de Bühlmann; sin embargo, la información del plan (la cartera) se parte en grupos de riesgo (grupos etarios (Schinzinger, Denuit & Christiansen, 2016, p. 72), por género (Tsai & Wu, 2020, p. 37) o por morbilidad) y se ponderan por las unidades expuestas de cada grupo. Luego se obtiene el factor de credibilidad “z” que a diferencia del modelo de Bühlmann éste depende de los promedios ponderados de los siniestros en función de las unidades expuestas que conforman cada partición. Para este modelo son necesario unos estimadores insesgados para obtener los años de experiencia, la heterogeneidad de la cartera y la variación de reclamaciones.

El modelo de Jewell se basa en el modelo anterior, con la salvedad de que se realiza una partición adicional a la información de la cartera, para obtener subcarteras; además, la misma se divide en subgrupos de riesgo (grupo etarios, por género o por morbilidad), por lo que este modelo requiere que la información de siniestralidad en cada una de las atenciones de un plan voluntario en salud se encuentre desglosada a un alto nivel de detalle, tal que el modelo incorpore mayor cantidad de información por lo cual es más probable que genere mejores resultados en porcentajes de siniestralidad, mayor rentabilidad en la cartera de riesgo, etc. Debido a la incorporación y segmentación de la cartera a este modelo se le conoce también como “jerárquico”. En este caso, se pondera por las unidades expuestas de cada grupo y modelo, luego se obtiene el factor de credibilidad “z” que es similar al modelo Bühlmann-Straub ya que depende de los promedios ponderados de los siniestros en función de las unidades expuestas que constituyen a cada grupo de riesgo que se tarifica.

Para el *cuarto de los modelos*, se referencia la metodología llamada ley de fallas de Weibull, que se emplea para resolver problemas de ingeniería de sistemas, hacer una adecuada programación, planeación y operación en el mantenimiento de equipos para determinar las fallas de sus componentes y pronosticar su comportamiento futuro (Mudholkar, Asubonteng & Hutson, 2009) (Xie, Tang & Goh, 2002). Para establecer el modelo es necesario transformar los datos para lograr la compatibilidad con un modelo habitual comprensible y bien conveniente (Jiang, 2013).

Este modelo también es conocido como la curva de la tasa de riesgo, se caracteriza por sus componentes físicos y su forma es llamada Curva de la Bañera. Esta curva presenta tres regiones claves para la interpretación en el tiempo. La región I conocida como mortalidad infantil, la región II conocida como el período de vida útil, por último, la región III representa el desgaste o fatiga (Roesch, 2012; Zhang, Xie, Tang & NG, 2005).

Finalmente, se trae a colación para la tarificación, el *quinto modelo*, la metodología de los derivados. Dentro de la técnica de las finanzas es una estrategia de cobertura identificar y gestionar los riesgos; esto permite que los agentes económicos puedan minimizar los riesgos a los que se encuentran sometidos en dicho mercado, sin pretender crear nuevos riesgos. Para la tarifa, el riesgo es todo lo que abarca el precio, el riesgo inherente al que está sometido.

El modelo de Black-Scholes se utiliza para valorar el precio de una opción financiera. En este caso se aplica a un derivado financiero para determinar la cobertura de la tarificación. Se empieza considerando una *opción*, en un derivado, lo que constituye un contrato que da a su tenedor el *derecho* a comprar (*opción call*) o que da el *derecho* de vender (*opción put*) un activo a un precio determinado, durante un período o en una fecha prefijada (Lamothe & Pérez, 2005). Las opciones funcionan de manera similar a un seguro, ya que permiten a su tenedor cubrirse ante fluctuaciones del precio del activo subyacente. Para tener este *derecho*, que en el sector asegurador es el suscriptor expuesto (o

asegurado), el tendedor del *derecho* debe pagar una prima, que se asemeja a la prima que cobran las compañías de seguro a sus suscriptores expuestos (o asegurados).

Los modelos propuestos, que se desarrollan para analizar las reclamaciones, tienen por objeto formar parte del estudio de un conjunto de opciones en los procedimientos actuariales en pro del cálculo de la tarificación, dispuesto a ser objeto de estudio para el sector. Se basa en un conjunto de métodos y técnicas científicamente sustentadas que se aplican al problema de estimar la tarifas óptima y suficiente. Por lo que este trabajo es una opción para los posibles procedimientos y metodologías con los que se determinan el valor tarifario óptimo y suficiente en la prestación de un plan voluntario de salud de acuerdo con la práctica actuarial actual (Yeo & Valdez, 2006, p. 624). El funcionamiento de la atención en todo siniestro viene asociado a un riesgo, siendo este siniestro para el presente documento una atención en un plan voluntario de salud, que se sustenta en un sistema de tarificación, que tiene un conjunto de principios técnicos en que se considera la estimación y el cálculo de una tarifa. A continuación, se presentarán las metodologías propuestas.

3. Metodologías

A continuación, se nombran las metodologías tradicionales actuariales utilizadas en la investigación con base en el riesgo individual, el riesgo colectivo y la teoría de la credibilidad. Se usa análisis de datos para determinar qué funciones de distribuciones se pueden ajustar a la historia y usar las metodologías actuariales, para compararlas con los métodos propuestos en el sector de la tecnología y en el sector financiero; siendo estos la tasa de fallas llamada también Curva de la Bañera y utilizando el modelo de opciones financieras de Black-Scholes. En adelante se tendrán algunas consideraciones de las variables que fueron necesarias para el uso de los modelos.

Dentro del contexto estadístico, las variables utilizadas en la evaluación de la tarificación contienen unas características que hacen esencial su identificación para su correcta aplicación en los ajustes de distribución. Para ello encontraremos que la base de datos contiene unas variables de tipo discretas y otras de tipo continuas, clasificación que es importante para agrupar los datos a una serie de distribuciones que ajustan dicha información y comportamiento. Las variables identificadas en los modelos son las *reclamaciones* de los siniestros atendidos, como también su correspondiente *costo*. Para alguno de los modelos, estas variables son transformadas en ratios conocidos como la *frecuencia* y la *severidad* (Jeong & Valdez, 2020, p. 23). Al igual, se tendrán en cuenta otras variables que servirán para dar más detalle a la base de datos como es la edad de los suscritos atendidos y no atendidos y el periodo en la que fue ocurrida dicha siniestralidad.

Todos los modelos adoptados en la tarificación deben incluir para sus estimaciones los parámetros que son básicos en cualquier metodología. Para ello la probabilidad de ocurrencia de un siniestro que si bien es la materialización del riesgo viene expresada en la unidad de tiempo frecuencia. Para el análisis de las variables y su distribución se evaluó su comportamiento segregado en cuanto a su ajuste y datos estadísticos resultantes. Lo anterior con el fin de realizar conjugación de estas distribuciones para homogenizar en una sola variable y mirar su resultado ajustado.

Tarifa

A continuación, se reseña la ecuación fundamental para los seguros establecida para el cálculo de la tarifa en un producto y con un contrato específico en una sociedad de libre mercado. De tal manera que, para definir dicho precio, al que el asegurador está dispuesto a vender y el consumidor está dispuesto a comprar para un periodo anual, y expresada en unidades monetarias (Werner & Modlin, 2019, p. 5), la *Tarifa (T)* debe reflejar los *Costos (S)*, los *Gastos (G)* y las *Utilidades* esperadas (*U*).

$$T = S + G + U \quad [1]$$

Así mismo, la expresión anterior debe asegurar un adecuado equilibrio en cada uno de los puntos, expresado en términos de tasa porcentual (Committee on Ratemaking Principles, 1988). La expresión contiene los porcentajes de los *gastos* (g) y la *utilidad* (u) con respecto a la *Tarifa* (T). Entonces la ecuación estará dada por:

$$T = S + g * T + u * T \quad \text{o} \quad T = \frac{S}{1-g-u} \quad [2]$$

Si bien la tarifa para un seguro se determina por la promesa de proporcionar una compensación en caso de que ocurra un evento de pérdida específica, por lo que los costos no se determinan en un punto de venta, y como resultado deben ser estimados, por lo que la tarifación es prospectiva para los seguros (OCDE, 2016; Werner & Modlin, 2019). A continuación, se utilizará el modelo de pérdidas agregadas para definir la tarifa para el futuro.

Modelo de Pérdidas Agregadas

Ahora, para estimar los costos, es necesario considerar las atenciones futuras de los siniestros inciertos, para ello se utiliza el modelo de pérdidas agregadas (Klugman, Panjer, & Willmot, 2019; Parodi, 2015). Se deben estimar dichos *Costos* (S) en un periodo anual y conocer la *cantidad de reclamaciones* (N) del periodo, además del *costo* (X_i) de cada una de las *i-ésimas* reclamaciones.

$$S = X_1 + X_2 + \dots + X_N \quad [3]$$

Para estimar la distribución de las pérdidas agregadas, es necesario estimar cuál sería el promedio de las pérdidas totales [$E(S)$] en el periodo de tiempo dado. Lo anterior se resuelve fácilmente con la media de la *cantidad de reclamaciones* de las pérdidas [$E(N)$] multiplicada por la media del *costo* de las pérdidas [$E(X)$]:

$$E(S) = E(N) * E(X) \quad [4]$$

Ahora, conociendo el promedio de las pérdidas agregadas, se puede encontrar la volatilidad anual de los costos, dado por la desviación estándar de (S). Luego la varianza $Var(S)$ está dada por:

$$Var(S) = E(N) * Var(X) + E(X)^2 * Var(N) \quad [5]$$

En las expresiones anteriores, se logra el equilibrio deseado a nivel agregado y la importancia a nivel de riesgo asociado de los costos medios transferidos de una cartera (Committee on Ratemaking Principles, 1988).

Prima pura

Las empresas del sector salud dependen de ratios para supervisar y evaluar la idoneidad de las tarifas, entre ellos la frecuencia, la severidad y la prima pura (Werner & Modlin, 2019, p. 7). La primera ratio utilizada para este artículo y que se utiliza en la actuaría es la *frecuencia* de los siniestros (r). La frecuencia representa el número de siniestros por unidad expuesta, y relaciona las ocasiones en las que se detecta la presencia de eventos causales de riesgo, siendo esta la razón entre la *cantidad de reclamaciones* atendidas (N) con respecto a la población total de los *expuestos suscritos* (e). Tal que:

$$Frecuencia = \frac{\text{cantidad de reclamaciones}}{\text{expuestos suscritos}} \quad \text{o} \quad r = \frac{N}{e} \quad [6]$$

La segunda ratio utilizada en este artículo y en la actuaría es la *severidad* de los siniestros (X), representada como el costo de pérdida por unidad de reclamación, es la razón entre a los *costos* (S) de la cartera con respecto a la *cantidad de reclamaciones* atendidas (N). Representa el costo promedio por

unidad de servicio de la siniestralidad. Siendo esta la variable que representa la importancia del impacto de los eventos de riesgo de enfermar sobre los resultados de la empresa.

$$Severidad = \frac{Costos}{cantidad\ de\ reclamaciones} \quad o \quad X = \frac{S}{N} \quad [7]$$

Finalmente, las ratios anteriores permiten calcular el costo puro de cada siniestro, el resultado es la *prima pura* (P), no es más que el producto entre las ratios *frecuencia* y *severidad de los siniestros*.

$$Prima\ pura = Frecuencia * Severidad \quad o \quad P = r * X \quad [8]$$

Utilizando la fórmula (6) y (7) en la ecuación (8):

$$Prima\ pura = \frac{Costos}{expuestos\ suscritos} \quad o \quad P = \frac{S}{e} \quad [9]$$

Ahora para estimar la prima pura media $E(P)$, será necesario estimar la frecuencia media $E(r)$ y la severidad media $E(X)$.

$$E(P) = E(r) * E(X) \quad o \quad E(P) = E\left(\frac{N}{e}\right) * E\left(\frac{S}{N}\right) \quad [10]$$

Las ecuaciones anteriores serán la base en la que se evaluará la distribución de probabilidad de ajuste a cada uno de los cinco métodos propuestos.

A continuación, se introducen la formulación a los métodos propuestos en objetivo a la construcción a un modelo de probabilidad del que describe las reclamaciones agregadas para un sistema de ocurrencia de seguros en un plan voluntario de salud en un periodo fijo del tiempo, derivando un conjunto de ecuaciones para la estimación de la tarifa. Además, se presentará el método de aprobación a los modelos y su prueba de hipótesis.

3.1. Método del riesgo individual

Este método trata de identificar por qué la mejor representación en el modelo está en seleccionar solo una variable para la modelación; ya sea solo la *frecuencia* o solo la *severidad*. Para el primer caso de modelación, se debe tomar la *frecuencia* como una variable aleatoria determinada, se relaciona en general por el riesgo subyacente del afiliado expuesto, particularmente sus razones se dan por la variabilidad que puede tener una mayor representación en diferente eventos; ya sea por cuestiones de crecimiento de la empresa, por una mayor cantidad de afiliados expuestos o por una circunstancia de evento masivo en la sociedad, como por ejemplo una epidemia local eleva la frecuencia de uso en el sector salud, como pasó en épocas anteriores. Para el segundo caso de modelación, donde la estimación se realiza para la *severidad* como una variable aleatoria determinada, en general se relaciona por la práctica de los ajustadores del siniestro, particularmente sus razones de variabilidad se pueden analizar por los eventos de aumentos considerables en el costo, o por aumento en crecimiento tecnológico de la empresa, o por eventos macroeconómicos que influyen en el aumento de costos en los insumos en la salud, entre otros.

3.1.1. Modelización de la severidad y reclamaciones constante

Para el modelo de riesgo individual con la modelización de la severidad como un variable aleatoria y las reclamaciones como una constante en el modelo se tiene que, la severidad son los costos individuales (X_{ij}) incurridos para todo riesgo de la cartera j –ésima $\in \{1, \dots, m\}$ y para cada una de las i –ésimas reclamaciones donde $i \in \{1, \dots, n\}$ y que ocurren durante un periodo de tiempo determinado tal que las pérdidas agregadas están dadas por $S = \sum_{i=1}^n X_{ij}$.

Utilizando el modelo de pérdidas agregadas anterior, se requieren unos supuestos adicionales: los costos (X_{ij}) son variables aleatorias independientes no negativas y no distribuidas de forma idéntica, por tanto, este modelo tiene la característica de ser un modelo heterogéneo. Usualmente, la distribución de los costos (X_{ij}) tiene una masa de probabilidad de cero, que corresponde a la probabilidad de no pérdidas o de que no haya riesgo de evento en la cartera j – *ésima* (Klugman, Panjer, & Willmot, 2019, p. 148). Mientras que la distribución de las reclamaciones (N_j) es obtenida de la distribución de los costos (X_{ij}) . Usando el enfoque de este modelo la severidad (X_{ij}) es modelada, mientras que los siniestros (N_j) será una constante en el modelo.

Ahora, los costos individuales (X_{ij}) al no ser necesariamente distribuidas idénticamente, como se acaba de mencionar, permiten que la esperanza de las pérdidas agregadas $E(S) = \sum_{i=1}^n E(X_{ij})$. Ahora, bajo el supuesto de independencia en los costos individuales (X_{ij}) implica que $COV(X_{1j}, X_{2j}) = 0$ para todo $(i_1 \neq i_2)$ de manera que la varianza de las pérdidas agregadas es $Var(S) = \sum_{i=1}^n Var(X_{ij})$. Además, siendo las reclamaciones (N_j) constantes que se calculan como un promedio aritmético de los datos históricos de las reclamaciones, se tiene que las pérdidas agregadas (S) en el modelo de riesgo individual para la modelización de la severidad, está dada por la media y la varianza respectivamente:

$$E(S) = N * E(X) \quad [11]$$

$$Var(S) = N * Var(X) \quad [12]$$

3.1.2. Modelización de las reclamaciones y severidad constante

Para el modelo de riesgo individual con la modelización de las reclamaciones como un variable aleatoria y la severidad como una variable constante en el modelo se tiene que, las reclamaciones individuales (N_{ij}) incurridas para todo riesgo de la cartera j – *ésima* $\in \{1, \dots, m\}$ y para cada una de las i – *ésimas* reclamaciones donde $i \in \{1, \dots, n\}$ y que ocurren durante un periodo de tiempo determinado tal que las pérdidas agregadas están dadas por $S = \sum_{i=1}^n N_{ij}$.

Al utilizar el modelo de pérdidas agregadas anterior, se tienen, además, los siguientes supuestos, las reclamaciones (N_{ij}) son unas variables aleatorias independientes no negativas y no distribuidas de forma idéntica, por lo que este modelo tiene la característica de ser un modelo heterogéneo. Las distribuciones de las reclamaciones (N_{ij}) tiene una función de masa de probabilidad, que asigna probabilidades de cero a la no ocurrencia de siniestros o a que no haya riesgo de evento en la cartera j – *ésima*, así como es reseñado por los autores (Klugman, Panjer, & Willmot, 2019, p. 148). Mientras que la distribución de los costos (X_j) es obtenida de la distribución de las reclamaciones (N_{ij}) . Usando el enfoque de este modelo donde las reclamaciones (N_{ij}) son modelados, mientras que la severidad (X_j) será una constante en el modelo.

Ahora, a través de las reclamaciones individuales (N_{ij}) se calcula la esperanza de las pérdidas agregadas como $E(S) = \sum_{i=1}^n E(N_{ij})$. El supuesto de independencia de las reclamaciones individuales (N_{ij}) implica que $COV(N_{1j}, N_{2j}) = 0$ para todo $(i_1 \neq i_2)$, por tanto, la varianza de las pérdidas agregadas está dada por $Var(S) = \sum_{i=1}^n Var(N_{ij})$. Además, como la severidad (X_j) es constante se calcula como el promedio aritmético de los datos históricos de los costos. En consecuencia, se tiene que la variable aleatoria pérdida agregada (S) para el modelo de riesgo individual para la modelización de los siniestros, tiene por media y varianza, respectivamente:

$$E(S) = X * E(N) \quad [13]$$

$$Var(S) = X * Var(N) \quad [14]$$

3.2. Método del riesgo colectivo

Este modelo considera la dependencia entre las variables aleatorias el número de reclamaciones y el costo de la reclamación, ambas involucradas en la suma del monto total de la reclamación (Cossette, Marceau, & Mtlai, 2019). Las series frecuencia y severidad tienen correlación cruzada (Lee & Shi, 2019, p. 115).

Las pérdidas agregadas están dadas por $S = \sum_{i=1}^N X_{ij}$.

Las variables aleatorias de los costos (X_{ij}) y las reclamaciones (N_j) tiene los siguientes supuestos: $(X_{1j}), (X_{2j}), \dots, (X_{nj})$ son independientes, no negativas de una cartera j y si distribuidas idénticamente, esto implica que $COV(X_{ij}, X_{kj}) \neq 0$. Usualmente, la distribución de cada (X_{ij}) tiene una función de masa de probabilidad, tal que, la probabilidad de que no haya evento en la cartera j es de cero (Klugman, Panjer, & Willmot, 2019, p. 148). Las reclamaciones (N_j) y la severidad (X_j), se modelan por separado y se aplica la convolución de ambas distribuciones.

Finalmente, para las pérdidas agregadas (S) en el modelo de riesgo colectivo, se estima su media y varianza:

$$E(S) = E(N) * E(X) \quad [15]$$

$$Var(S) = E(X)^2 * Var(N) + E(N) * Var(X) \quad [16]$$

3.3. Método de la Credibilidad

Como se mencionó anteriormente, el objetivo de estimar la tarifa o el precio de un plan es evaluar correctamente los costos potenciales. Cuando los suscriptores expuestos es lo suficientemente grande como para tener una experiencia creíble en reclamaciones, esos datos de reclamaciones se pueden utilizar para modificar la tarifa que se habría cobrado solo por un plan de clasificación. Si el plan de clasificación es relativamente simple, la credibilidad de la experiencia de reclamo individual será mayor; si el plan de clasificación es más preciso, la experiencia de reclamación de riesgo individual tendrá menos credibilidad. En un plan de clase simple, hay una mayor cantidad de suscriptores que recibirán las mismas tarifas. Por lo general, se deduce que habrá más heterogeneidad de costos dentro de cada grupo. Por lo tanto, si un suscriptor ha tenido o no un reclamo, será más indicativo de si el costo real (desconocido) de ese individuo difiere del costo promedio del grupo. Si todos los suscriptores dentro del grupo tuvieran exactamente el mismo costo, la experiencia real de las reclamaciones sería puramente fortuita y no tendría credibilidad.

Para determinar la tarificación con base en el concepto de credibilidad, es indispensable la experiencia en las reclamaciones. Para un producto nuevo, no existe esta credibilidad ya que carecería de historia y de experiencia, no habría conocimiento de las reclamaciones.

A lo anterior, la importancia en la aplicación de esta metodología, a tal punto de encontrar un equilibrio en la información que tiene el suscriptor y la información que posee todo un grupo de suscriptores. Ya que, de acuerdo a las características particulares de cada cartera, es muy poco factible aplicar tarifas iguales para riesgos distintos. Es por esto que para la aplicación de este modelo en la presente investigación se tomará una cartera en específico (Tsai & Wu, 2020). A continuación, se presentan la estructura del modelo de credibilidad.

Se tiene el costo estimado (X_{ij}) y representa el costo de las experiencias de las reclamaciones para cada grupo etario representadas en las i –ésimas reclamaciones del grupo, tal que $i \in \{1, \dots, n\}$, para los individuos de la cartera j –ésima en un periodo (t). Sea el costo total estimado de toda la cartera (X_T). Sea (S_{ij}) el costo estimado en un período futuro para el modelo credibilidad. Para aplicar el modelo de la credibilidad, se utiliza la experiencia de pérdida en un período histórico,

con el fin de predecir con mayor precisión la experiencia de pérdida en el siguiente período, con la siguiente expresión:

$$S_{ij} = Z * X_{ij} + (1 - Z) * X_T \quad [17]$$

Luego, la fórmula de la teoría general de la credibilidad está dada por un estimador, el factor de credibilidad Z . A continuación, es necesario determinar la varianza de las medidas hipotéticas de la estructura $VHM = Var[E(s_{Fi}|\theta)] = \tau^2$ y el valor esperado de la varianza del proceso $EPV = Var[E(s_{Fi}|\theta)] = \sigma^2$, tal que:

$$Z = \frac{Var[E(X_{Fi}|\theta)]}{Var(X_{Fi})} = \frac{Var[E(s_{Fi}|\theta)]}{Var[E(s_{Fi}|\theta)] + E[Var(X_{Fi}|\theta)]} \quad [18]$$

Tal que:

$$Z = \frac{\tau^2}{\tau^2 + \sigma^2} \quad [19]$$

Para lo que el denominador es considerado la varianza total.

3.4. Método Tasa de Fallas de Weibull (Curva de la Bañera)

La curva de la tasa de riesgo se caracteriza por sus componentes físicos, en su forma es llamada Curva de la Bañera. Esta curva presenta tres regiones claves para la interpretación en el tiempo.

La región I conocida como mortalidad infantil, representando el período de arranque donde la tasa de riesgo (S) desciende a medida que pasa el tiempo. Roesch (2012) centra su atención en esta zona para mostrar la relación de calidad y confiabilidad, aspectos que utiliza para mejora y reducción de riesgos del cliente. La región II es conocida como el período de vida útil, siendo ésta la representación de la fase de operación normal que se caracteriza por tener una tasa de riesgo (S) constante, donde las fallas ocurren al azar y se aplica la distribución exponencial. Para esta región, para estimar un buen modelo de la curva de la bañera su parte plana debe tener una extensión relativamente ancha (Zhang, Xie, Tang, & NG, 2005). Por último, la región III representa el desgaste o fatiga y se caracteriza por un rápido incremento en la tasa de riesgo (S) en el tiempo.

Para este método se determina la tasa de fallas o tasa de riesgo (S_{it}), como medida de la velocidad de ocurrencia y viene dada por el número de fallas, si bien para este artículo será el costo (X_{ij}) la representación de la velocidad monetaria de las reclamaciones por unidad de tiempo t , a razón del número de componentes expuestos a la falla (e).

Se representa el costo de las experiencias de las reclamaciones para cada grupo etéreo talque las i – ésimas son reclamaciones del grupo, para $i \in \{1, \dots, n\}$, pertenecen a todo riesgo de la cartera j – ésima $\in \{1, \dots, m\}$.

$$S_{ij} = \frac{X_{ij}}{e} \quad [20]$$

3.5. Método de Black Scholes de opciones financieras

Sea el tenedor el que tiene el derecho a hacer uso de la opción cuando el precio del activo subyacente (S_t) en el tiempo (t), es superior al precio de ejercicio o precio pactado (S_E). Por el contrario, si el precio del activo subyacente (S_t) en el tiempo (t), es igual o inferior al precio de ejercicio o pactado (S_E), su tenedor no ejercerá la opción y la opción vencerá sin valor alguno (Marín & Rubio, 2001); en resumen, los pagos de la *call europea* al vencimiento pueden resumirse como:

$$\begin{cases} S_t - S_k, & \text{si } S_t > S_E \\ 0, & \text{si } S_t < S_E \end{cases} \quad [21]$$

El valor de una opción se denomina prima y se asemeja a la prima de riesgo calculada por los métodos tradicionales actuariales; la fórmula anterior más popularizada en tiempo continuo para valorar opciones (Black & Scholes, 1973).

Luego, el precio del activo es modelado por el siguiente movimiento browniano geométrico:

$$dS = \mu * S_t * dt + \sigma * S_t * dW \quad [22]$$

donde (μ) es la tasa de retorno esperada, (σ) es la volatilidad del retorno del activo y (dW) es un movimiento browniano (Black & Scholes, 1973), haciendo uso del cálculo estocástico de Ito y de las ecuaciones diferenciales estocásticas parciales, deriva las siguientes fórmulas para valorar el precio teórico de una opción *call* y una opción *put* europea cuando la acción no paga dividendos (Shreve, 2004).

$$\text{CALL} \quad c = S_t * N(d_1) - S_E * e^{-rt} * N(d_2) \quad [23]$$

$$\text{PUT} \quad p = S_E * e^{-rt} * N(-d_2) - S_t * N(-d_1) \quad [24]$$

donde:

$$d_1 = \frac{\ln\left(\frac{S_t}{S_E}\right) + \left(r + \frac{\sigma^2}{2}\right) * T}{\sigma * \sqrt{T}} \quad [25]$$

$$d_2 = \frac{\ln\left(\frac{S_t}{S_E}\right) + \left(r - \frac{\sigma^2}{2}\right) * T}{\sigma * \sqrt{T}} \quad [26]$$

Las variables necesarias para la aplicación del modelo de Black & Scholes (1973) son: (S_t), precio del activo subyacente en el momento de la valoración de la opción; (S_E), el precio de ejercicio como el precio pactado al iniciar el contrato; (r), la tasa de interés en tiempo continuo; (T), plazo de ejercicio en años; (σ), la volatilidad de la variación del precio del subyacente en términos anuales; (c), la prima de la opción de compra o (prima del mercado); y (p), la prima de la opción de venta (Lamothe & Pérez, 2005) o (prima estimada).

Para el ejercicio a manera de práctica se utilizará la prima de la opción de compra (c), que no es más que la prima del mercado de la competencia.

4. Metodología para la selección del modelo

En este apartado del documento se presentan las diferentes metodologías para determinar el modelo probabilístico de donde provienen los datos.

Para este procedimiento de inferencia de los datos (Lenhard, 2006), utilizando criterios probabilísticos, se plante la siguiente Prueba de Hipótesis de investigación para realizar el análisis lógico por un algoritmo de decisión donde el objetivo del análisis es aceptar o rechazar la hipótesis nula (H_0) bajo el criterio de decisión acerca de su validez en la función de distribución, con la prueba de significación determinado por el valor P (Monterrey, 2012). Para esto, los resultados son apoyados computacionalmente con @ Risk mediante la técnica estadística paramétrica.

Es importante para la selección del modelo, la elección de las variables de clasificación para la estimación de la tarificación; estadísticamente se debe basar en cuatro criterios de tipo social, legal, operacional y estadístico (Finger, 2001). Principalmente para el criterio estadístico debe tener tres características que deben ser significativas, homogéneas y creíbles. Y, por último, para la función de distribución de probabilidad (f) de la variable analizada para el cálculo de la prima pura es evaluada bajo tres criterios arrojados por valores en @ Risk:

1. *La mejor explicación* de las distribuciones estimadas evaluadas por el criterio de ajuste de bondad con el menor valor en cada uno de los tres estadísticos.
2. *La validación del ajuste* de las distribuciones estimadas aceptadas (H_0) estadísticamente significativas para el modelo con el criterio el Valor del estadístico (Estadístico de prueba) vs Valor crítico (grado de significancia α del 5%) valorada bajo la siguiente prueba de hipótesis a un nivel de confianza $1 - \alpha$ del 95%.
3. *La selección del modelo* de las distribuciones estimadas evaluadas y estimadas a un nivel de confianza $1 - \alpha$ del 95%, se determinan bajo el criterio del P-valor. La distribución de mayor P-valor que cumplió con los dos criterios anteriores (mejor explicación y validación del ajuste), es la función de distribución de probabilidad (f) seleccionada para las variables en cada metodología evaluada.

4.1. Prueba de bondad

A continuación, se presentan los estadísticos evaluadores para la selección de la función de distribución de probabilidad (f) que mejor explica el comportamiento a las variables que servirá para los estadísticos del modelo.

4.1.1. Mejor explicación del modelo y sus estadísticos de Ajuste de Bondad

Los estadísticos de ajuste de bondad prueban el nivel de coincidencia entre la muestra de los datos históricos de entrada y los datos de estimación ajustados para el futuro. El estadístico también contiene un nivel de confianza en el que los datos de entrada han sido producidos por la función de distribución. Por cada uno de estos estadísticos de ajuste, cuanto menor sea el valor estadístico, mejor es el ajuste.

Los estadísticos de ajuste en @ Risk evaluadores de los datos históricos de entrada el Criterio de Información Akaike (AIC), el Criterio de Información Bayesiano (BIC), Chi Cuadrada (CHI). Además, los estadísticos de ajuste Anderson Darling (AD) y Kolmogorov Smirnov (KS).

Tabla 1. Pruebas de Bondad de Ajuste.

Ajuste de bondad	Fórmula del estadístico	Parámetros	
Chi-Cuadrado (CHI)	$\chi^2 = \sum \frac{(o_i - e_i)^2}{e_i}$	o_i : Representa cada frecuencia (severidad) observada.	e_i : Representa cada frecuencia (severidad) esperada o teórica.
Anderson Darling (AD)	$A^2 = -N - S$	N : Es el número de observaciones.	
	$S: \sum_{k=1}^N \frac{2k-1}{N} [\ln F(Y_k) + \ln(1 - F(Y_{N+1-k}))]$	$F(Y)$: Es la distribución de probabilidades acumulada.	Y_k : Son los datos obtenidos en la muestra, ordenados de menor a mayor.
Criterio Información Akaike (AIC)	$AIC = -2k - 2 * \ln(L)$	k : Es el número de parámetros del modelo.	$\ln(L)$: Es la función de log-verosimilitud para la distribución de probabilidad.

Fuente: Elaboración propia.

Para el presente documento los datos históricos se evalúan con los estadísticos de ajuste CHI y AD para determinar la *mejor explicación y validación del ajuste* según la prueba de hipótesis. El estadístico AIC no es una prueba del modelo evaluadora de una hipótesis, este estadístico proporciona un medio comparable entre dos modelos como una herramienta para determinar la *selección del modelo*. El menor estadístico AIC servirá para elegir la mejor distribución.

Ahora, ya nombrado los ajustes de bondad, se debe tener en cuenta para calcular las mejores distribuciones de los datos históricos se utilizan dos métodos en @ Risk:

1. Para los *datos históricos de entrada*, los parámetros de la distribución de probabilidad (f) se proyectan utilizando Estimadores de Máxima Probabilidad (Maximum Likelihood Estimators - MLE) o ligeras modificaciones de MLE.
2. Para los *datos históricos de entrada*, la estimación de densidad y acumulativos (F), llamados colectivamente datos de curva, se utiliza el método de mínimos cuadrados para minimizar el Error de Raíz Cuadrada de la Media (RMSErr) que hay entre los puntos de la curva y la función observada (teórica).

4.1.1.1. *Ajuste de Bondad Chi-Cuadrada χ^2 (CHI)*

En este estadístico, para determinar el valor crítico en donde inicia la región de rechazo es necesario determinar el nivel de confianza, el número de clases y los parámetros que definen la distribución de probabilidad; ya que de estos dependen los grados de libertad de la prueba. Dado al momento de realizar el ajuste de datos se pueden encontrar diferentes distribuciones de probabilidad, y estas en algunos casos pueden tener un solo parámetro, como la distribución exponencial λ , o como en la mayoría de los casos tienen dos, como en la distribución Normal (μ ; σ) o en algunas oportunidades tres como en la distribución Triangular (a ; c ; b).

Para evitar tener en cuenta el número de clases y el número de parámetros según la distribución de probabilidad a ajustar, se va a tomar para la determinación de la región de rechazo, como base 300 grados de libertad (donde los grados de libertad son mucho menores que los 365 posibles menos los parámetros, y es un valor que aparece en las tablas χ^2) para lo cual el estadístico Chi queda definido como $\chi^2_{0,95,300} > 260,8781$, es decir, valores por encima de 260,8781 entran en la región de rechazo, por lo cual se rechazaría la hipótesis nula en favor de la hipótesis alternativa.

4.1.1.2. *Ajuste de Bondad Anderson Darling (AD)*

Al igual que la prueba Chi-cuadrada χ^2 , la AD siendo una prueba no paramétrica mide qué tan bien siguen los datos una distribución específica. Para un conjunto de datos y distribución en particular, mientras mejor se ajuste la distribución a los datos, menor será este estadístico. Esta prueba es una variante a la prueba de ajuste de Kolmogorv-Smirnov y será utilizada en vez de ésta, ya que AD es considerada una alternativa más potente, ya que en ésta se quiere dar más peso a las colas de la distribución, siendo ésta ideal para colas pesadas.

Aunque la prueba de AD puede ser aplicada a cualquier distribución, no se dispone de tablas para todos los casos. A continuación, se presentan las regiones de rechazo para algunas distribuciones con un nivel de confianza del 95%: Distribución Normal 0,751, Distribución Exponencial 1,326, Distribución Weibull 0,757, Distribución Log logística 0,660. Dado a la imposibilidad de que no se cuenta con regiones de rechazo para todas las distribuciones, se tomará como referencia la región de rechazo de la Loglogística 0,660 para las distribuciones de probabilidad que no se tiene dato en una tabla, que es la región de rechazo más restrictiva de las regiones conocidas y con base en ésta se aceptará o se rechazará.

4.1.1.3. Ajuste de Bondad Criterio de Información de Akaike (AIC)

Es una medida de la calidad relativa del ajuste de datos de una distribución de probabilidad. El AIC proporciona un medio para la *selección del modelo*, pero no de acuerdo a una prueba de hipótesis, sino que entre varias posibles distribuciones de probabilidad se debe escoger la que tiene el menor valor AIC. A diferencia de la prueba Chi-cuadrado χ^2 , este test trata de penalizar un exceso de parámetros ajustados.

Como se mencionó anteriormente el AIC clasifica las distribuciones de probabilidad de acuerdo al valor obtenido del AIC y mientras menor sea este, mejor es el ajuste. Por si solo no dice mucho, ya que, si todas las distribuciones de probabilidad no son apropiadas para determinar la *mejor explicación y validación del ajuste*, él no es capaz de identificar este hecho, por eso debe ir acompañado de otras pruebas de ajuste, tales como la Chi-cuadrado χ^2 y la Anderson Darling.

A continuación, se evalúan los estadísticos de ajuste, se usan para dos propósitos relacionados pero diferentes. *La validación del ajuste* es el proceso por el que se determina si una distribución ajustada es un “aceptable” ajuste para los datos de entrada, mientras que *la selección del modelo* es el proceso de elección de un tipo determinado de distribución mejor ajustada sobre otro.

4.1.2. Validación del ajuste

Para la validación del ajuste se usa la técnica de carácter estocástica Bootstrap paramétrico (Efron & Gong, 1983) ya que se parte de una determinada función de distribución en @ Risk que ayuda a superar limitaciones en los análisis en los que uno o más parámetros eran ajustables en la función de distribución de probabilidad que en muchos casos resultaban inaceptables en el modelo. Esta técnica evita los errores de Tipo II denominado falso negativo, es el error que se comete al aceptar la Hipótesis nula H_0 cuando en realidad es falsa y debería ser rechazada en la población. Esta situación experimental es relacionada con el grado de significancia estadístico α que para los modelos de este documento se evalúa al 5%.

Como ya se había nombrado para el presente documento los estadísticos de ajuste CHI y AD son los que determinan la validación del ajuste según la prueba de hipótesis. En caso de que uno de los dos sea rechazado, con al menos uno de los dos estadísticos sea aceptado, se seleccionara como óptimo la función de distribución de probabilidad f . Lo anterior, con el fin de evitar errores de Tipo I denominado falso positivo, es el error que se comete al rechazar la Hipótesis nula H_0 cuando en realidad es verdadera y debería ser aceptada en la población.

4.1.2.1. Prueba de Hipótesis

H_0 . Hipótesis Nula: Los datos se ajustan o siguen una distribución específica.

H_a . Hipótesis Alternativa: Los datos no se ajustan o no siguen una distribución específica.

4.1.2.2. Región de rechazo

Si el Valor del estadístico es menor al Valor crítico se acepta H_0 . Se rechaza H_a .

Si el Valor del estadístico es mayor al Valor crítico se acepta H_a . Se rechaza H_0 .

En la validación del ajuste @ Risk arroja resultado de datos “N/D” para algunas funciones de distribución de probabilidad f estimada reportando no existencia de dependencia entre *datos históricos de entrada* y *datos de estimación ajustados* por lo que estas funciones no entran en la validación del análisis y por ende se rechazan, o se evalúa si se genera una distribución manual, la distribución triangular, o finalmente, la validación del ajuste se aprueba con el criterio de AIC y será el criterio para la selección del modelo.

4.1.3. Selección del modelo

Para la *selección del modelo* de las distribuciones estimadas evaluadas y estimadas a un nivel de confianza $1 - \alpha$ del 95%, se determinan bajo el criterio del P-Valor. La distribución de mayor P-Valor que cumpla con los dos criterios anteriores de la *mejor explicación* y de la *validación del ajuste*, es la función de distribución de probabilidad seleccionada para las variables en cada metodología evaluada.

Con el criterio P-Valor, se concluye si cada función es o no significativa para el modelo a pesar de que fueran las mejores distribuciones explicadas arrojadas por @ Risk.

5. Aplicación de modelos y Resultados

Para calcular la función distribución de probabilidad en las variables analizadas se realizaron 1000 remuestreos en @RISK. A su vez para la variable costos (S), se utilizó el método de simulación Montecarlo con 1000 iteraciones para estimar la esperanza de la variable para un periodo (t). Se obtiene la estimación de la variable costos (S^*) a continuación se describe el procedimiento realizado.

5.1. Modelo de Riesgo Individual

- a. Se tomó la variable aleatoria de las reclamaciones diarias del plan para el año (t) para la población que cumplía las características de homogeneidad.
- b. Se estableció la mejor distribución de los datos para la variable reclamaciones con el algoritmo del software @RISK.
- c. Se realizó la convolución entre la FDP de las reclamaciones y la severidad constante en este modelo.
- d. Una vez obtenida la convolución, se generaron los números aleatorios con la simulación.
- e. El valor de (S^*) corresponde a las 1000 réplicas de este procedimiento.

Luego:

- f. Se tomó la variable aleatoria severidad diaria de las reclamaciones del plan para el año (t) para la población que cumplía las características de homogeneidad.
- g. Se estableció la mejor distribución de los datos para la variable severidad con el algoritmo del software @RISK.
- h. Se realizó la convolución entre la FDP de la severidad y las reclamaciones como constante en este modelo.
- i. Una vez obtenida la convolución, se generaron los números aleatorios con la simulación.
- j. El valor de (S^*) corresponde a las 1000 réplicas de este procedimiento.

Luego se procede con:

- k. Los dos valores anteriores de (S^*) se promedian, y finalmente se multiplican con la población expuesta total del año (t), para hallar el Modelo de Pérdidas Agregadas para el Modelo de Riesgo Individual.

1. Las pérdidas agregadas se dividen por la población expuesta total del año (t) para obtener la prima pura.

5.2. Modelo de Riesgo Colectivo

- a. Se tomaron las variables aleatorias reclamaciones y la severidad diaria de las reclamaciones del plan para el año (t) para la población que cumplía las características de homogeneidad.
- b. Se estableció la mejor distribución de los datos para cada una de las variables con el algoritmo del software @RISK.
- c. Se estableció la mejor distribución de los datos para la variable pérdidas en el grupo etéreo uno (1) con el algoritmo del software @RISK.
- d. Una vez obtenida la convolución, se generó los números aleatorios con la simulación.
- e. El valor de (S^*) corresponde a las 1000 réplicas de este procedimiento.

Luego se procede con:

- f. Se multiplican (S^*) con la población expuesta total del año (t), para hallar el Modelo de Pérdidas Agregadas para el Modelo de Riesgo Colectivo.
- g. Las pérdidas agregadas se dividen por la población expuesta total del año (t) para obtener la prima pura.

Tabla 2. Prima del Modelo de Riesgo Individual.

1	Modelo de riesgo	FDP frecuencia	FDP Severidad	X ² Frecuencia	X ² Severidad	FDP frecuencia	FDP Severidad	Prima pura	Media	Modelo pérdidas agregadas	PRIMA PURA
Expuestos (e)	Individual	Función de distribución		Valor p – Prueba de hipótesis		Simulada	Simulada	Convolución	E[S]	LDA	
357843	θ_i Frecuencia	N _i Negbin	X _i Severidad constante	0,30 aceptada	2,511	72,697	186,193	186,269	186,269	66649755653.2	186254.4
357843	θ_i Severidad	N _i Frecuencia constante	X _i Normal		2,512	72,682	186,231	186,240	186,240		

Fuente: Elaboración propia.

Tabla 3. Prima del Modelo de Riesgo Colectivo.

2	Modelo De Riesgo	FDP Frecuencia	FDP Severidad	X ² Frecuencia	X ² Severidad	FDP Frecuencia	FDP Severidad	Prima pura	Media	Modelo pérdidas agregadas	PRIMA PURA
Expuestos (e)	Colectivo	Función de distribución		Valor p – Prueba hipótesis		Simulada	Simulada	Convolución	E[S]	LDA	
357843	θ_i Frecuencia-severidad	N _i Negbin	X _i Normal	0,30 aceptada	0,54 aceptada	2,511	72,682	186,156	186,134	66,606,614,151	186,134

Fuente: Elaboración propia.

5.3. Modelo de Riesgo de Credibilidad

- a. Se evaluaron nueve niveles de riesgo θ_i para este modelo; los nueve grupos etáreos que corresponde a las edades de 0 a 10 años, de 10 a 20 años, de 20 a 30 años, de 30 a 40 años, de 40 a 50 años, de 50 a 60 años, de 60 a 70 años, de 70 a 80 años y mayores de 80 años, cumpliendo así el principio de homogeneidad.
- b. Se tomó la variable aleatoria pérdidas de cada expuesto del plan en el año (t) , para el grupo etáreo uno (1) que comprende los expuestos de 0 a 10 años de edad.
- c. Se estableció la mejor distribución de los datos para la variable pérdidas del género femenino con el algoritmo del software @RISK.
- d. Se tomó la variable aleatoria pérdidas de la población total para cada expuesto del plan en el año (t) .
- e. Se estableció la mejor distribución de los datos para la variable pérdidas de la población total con el algoritmo del software @RISK.
- f. Se realizó el cálculo del factor Z de credibilidad.
- g. Se realizó la convolución entre la FDP de la variable aleatoria pérdidas de cada uno de los nueve grupos etáreos y la FDP variable aleatoria pérdidas de la población total con su respectivo valor de credibilidad.
- h. Una vez obtenida la convolución, se generaron los números aleatorios con la simulación.
- i. El valor de (S^*) corresponde a las 1000 réplicas de este procedimiento.

Luego:

- j. Se repite el procedimiento anterior para la variable aleatoria pérdidas de cada uno de los grupos etáreos restante.
- k. Los nueve valores anteriores de (S^*) se multiplican por su población expuesta e correspondiente para cada grupo etáreo en el año (t) . Finalmente se suman para hallar el Modelo de Pérdidas Agregadas para el Modelo de Riesgo de Credibilidad.
- l. Las pérdidas agregadas se dividen por la población expuesta total del año (t) para obtener la prima pura.

5.4. Modelo de Riesgo de Fallas

- a. Se evaluaron nueve niveles de riesgo θ_i para este modelo; los nueve grupos etáreos que corresponde a las edades de 0 a 10 años, de 10 a 20 años, de 20 a 30 años, de 30 a 40 años, de 40 a 50 años, de 50 a 60 años, de 60 a 70 años, de 70 a 80 años y mayores de 80 años, cumpliendo así el principio de homogeneidad.
- b. Se tomó la variable aleatoria pérdidas de cada expuesto del plan en el año (t) , para el grupo etáreo uno (1) que comprende los expuestos de 0 a 10 años de edad.
- c. Se estableció la mejor distribución de los datos para la variable pérdidas en el grupo etáreo uno (1) con el algoritmo del software @RISK.

- d. Se realizó la convolución entre la FDP de la variable aleatoria pérdidas del grupo etéreo uno (1), con respecto a la proporción de la población total con los expuestos correspondientes al grupo etéreo.
- e. Una vez obtenida la convolución, se generaron los números aleatorios con la simulación.
- f. El valor de la estimación (S^*) corresponde a las 1000 réplicas de este procedimiento en el grupo etéreo uno (1).

Luego:

- g. Se repite el procedimiento anterior para la variable aleatoria pérdidas de cada uno de los grupos etéreos restante.
- h. Los nueve valores anteriores de (S^*) se multiplican por su población expuesta e correspondiente para cada grupo etéreo en el año (t). Finalmente se suman para hallar el Modelo de Pérdidas Agregadas para el Modelo de Riesgo de Fallas.
- i. Las pérdidas agregadas se dividen por la población expuesta total del año (t) para obtener la prima pura.

Tabla 4. Prima del Modelo de Riesgo de Credibilidad.

3	Modelo De Riesgo	FDP Pérdidas Grupo etéreo i	FDP Pérdidas Grupo etéreo Total	Factor	Criterio Akaike Grupo etéreo i		Criterio Akaike Grupo etéreo total		FDP Pérdidas Grupo etéreo i	FDP Pérdidas Grupo etéreo total	Prima Pura	Media	Modelo Pérdidas Agregadas	PRIMA PURA
Expuestos (e)	Credibilidad	Función de distribución		Z	Valor criterio		Valor criterio		FDP Simulada	FDP Simulada	Convolución	E(S)	LDA	
37316	θ_i Grupo etéreo 1	S_{Ti} Invgauss	S_{Ti} Invgauss	0.4	9.83E+05	Aceptada	6.78E+06	Aceptada	292,487	260,659	200,308	200,103	77781416066	217362.1
24222	θ_i Grupo etéreo 2	S_{Ti} Invgauss	S_{Ti} Invgauss	0.24	4.64E+05	Aceptada	6.78E+06	Aceptada	208,002	260,659	196,451	196,292		
22061	θ_i Grupo etéreo 3	S_{Ti} Invgauss	S_{Ti} Invgauss	0.23	4.14E+05	Aceptada	6.78E+06	Aceptada	220,370	260,659	182,399	182,448		
25206	θ_i Grupo etéreo 4	S_{Ti} Invgauss	S_{Ti} Invgauss	0.23	4.39E+05	Aceptada	6.78E+06	Aceptada	216,624	260,659	189,109	190,547		
23608	θ_i Grupo etéreo 5	S_{Ti} Invgauss	S_{Ti} Invgauss	0.21	3.80E+05	Aceptada	6.78E+06	Aceptada	210,464	260,659	199,405	199,345		
36439	θ_i Grupo etéreo 6	S_{Ti} Invgauss	S_{Ti} Invgauss	0.27	5.73E+05	Aceptada	6.78E+06	Aceptada	213,518	260,659	201,788	202,285		
59721	θ_i Grupo etéreo 7	S_{Ti} Invgauss	S_{Ti} Invgauss	0.34	9.38E+05	Aceptada	6.78E+06	Aceptada	216,324	260,659	207,066	206,564		
64610	θ_i Grupo etéreo 8	S_{Ti} Invgauss	S_{Ti} Invgauss	0.4	1.13E+06	Aceptada	6.78E+06	Aceptada	252,492	260,659	222,778	222,608		
64660	θ_i Grupo etéreo 9	S_{Ti} Invgauss	S_{Ti} Invgauss	0.53	1.45E+06	Aceptada	6.78E+06	Aceptada	371,789	260,659	275,973	277,387		

Fuente: Elaboración propia.

Tabla 5. Prima del Modelo de Riesgo de Fallas.

4	Modelo de Riesgo		Criterio Akaike Grupo etéreo i		FDP Pérdidas grupo etéreo i	Prima pura	Media	Modelo Pérdidas Agregadas	PRIMA PURA
Expuestos (e)	Fallas (Curva Bañera)	Función de distribución	Valor criterio		FDP Simulada	Convolución	E[S]	LDA	
37316	θ_i Grupo etéreo 1	S_{GEi} Invgauss	9.83E+05	Aceptada	292,487	214,292	215,523	75887068389.7432	212068.300208569
24222	θ_i Grupo etéreo 2	S_{GEi} Invgauss	4.64E+05	Aceptada	208,002	164,610	164,754		
22061	θ_i Grupo etéreo 3	S_{GEi} Invgauss	4.14E+05	Aceptada	220,370	159,865	159,609		
25206	θ_i Grupo etéreo 4	S_{GEi} Invgauss	4.39E+05	Aceptada	216,624	163,629	163,748		
23608	θ_i Grupo etéreo 5	S_{GEi} Invgauss	3.80E+05	Aceptada	210,464	167,959	167,797		
36439	θ_i Grupo etéreo 6	S_{GEi} Invgauss	5.73E+05	Aceptada	213,518	173,634	173,291		
59721	θ_i Grupo etéreo 7	S_{GEi} Invgauss	9.38E+05	Aceptada	216,324	182,460	182,522		
64610	θ_i Grupo etéreo 8	S_{GEi} Invgauss	1.13E+06	Aceptada	252,492	218,556	220,347		
64660	θ_i Grupo etéreo 9	S_{GEi} Invgauss	1.45E+06	Aceptada	371,789	321,078	321,567		

Fuente: Elaboración propia.

5.5. Modelo de Riesgo de Black-Scholes

- a. Se evaluaron doce niveles de riesgo θ_i para este modelo; corresponde a los doce meses del año (t), cumpliendo así el principio de homogeneidad.
- b. Se tomó la variable aleatoria pérdidas de cada expuesto del plan en el año (t), para el mes uno (1).
- c. Se estableció la mejor distribución de los datos para la variable pérdidas para el mes uno (1) con el algoritmo del software @RISK.
- d. Se realizó la convolución entre la FDP de la variable aleatoria pérdidas del mes uno (1), con respecto a la proporción de la población total con los expuestos correspondientes a dicho periodo.
- e. Una vez obtenida la convolución, se generaron los números aleatorios con la simulación.
- f. El valor de la estimación (S^*) corresponde a las 1000 réplicas de este procedimiento del mes uno (1).
- g. El valor anterior corresponde al costo estimado (S_t^*) para el mes uno (1). También se puede identificar como el valor del Subyacente.
- h. El costo de ejercicio (S_E^*), es el valor pactado inicialmente en el contrato. Para el presente documento, utilizaremos el costo de ejercicio bajo el supuesto del valor promedio ponderado de la media de los costos (S) de cada mes y su respectivo peso de los expuestos de cada mes. Este costo de ejercicio es igual para todos los meses.
- i. La tasa de retorno r , es un 15% anual ya que es el retorno máximo exigido por la ley en Colombia. Esta tasa de retorno es igual para todos los meses.
- j. El tiempo T , es de 1 año ya que los datos se están evaluando para este periodo de tiempo. El tiempo es igual para todos los meses.
- k. La volatilidad σ para el mes uno (1), corresponde a la desviación porcentual de las pérdidas del mes uno (1).
- l. Se procede a calcular el valor de la opción call de Black-Sholes.
- m. Se procede a evaluar el criterio de la opción call de Black-Sholes.

Luego:

- n. Se repite el procedimiento anterior para la variable aleatoria pérdidas de cada uno de los meses restantes.
- o. Los doce valores anteriores de (S^*) se multiplican por su población expuesta e correspondiente para cada grupo etáreo en el año t .
- p. De los doce meses se evalúa aquellos costos generados por la opción call al no ejercerla, y se multiplican por su población expuesta e correspondiente para cada mes del año t .
- q. Finalmente se suman los dos resultados anteriores para hallar el Modelo de Pérdidas Agregadas para el Modelo de Riesgo Black-Scholes.
- r. Las pérdidas agregadas se dividen por la población expuesta total del año t para obtener la prima pura.

Tabla 6. Prima del Modelo de Riesgo Black-Scholes.

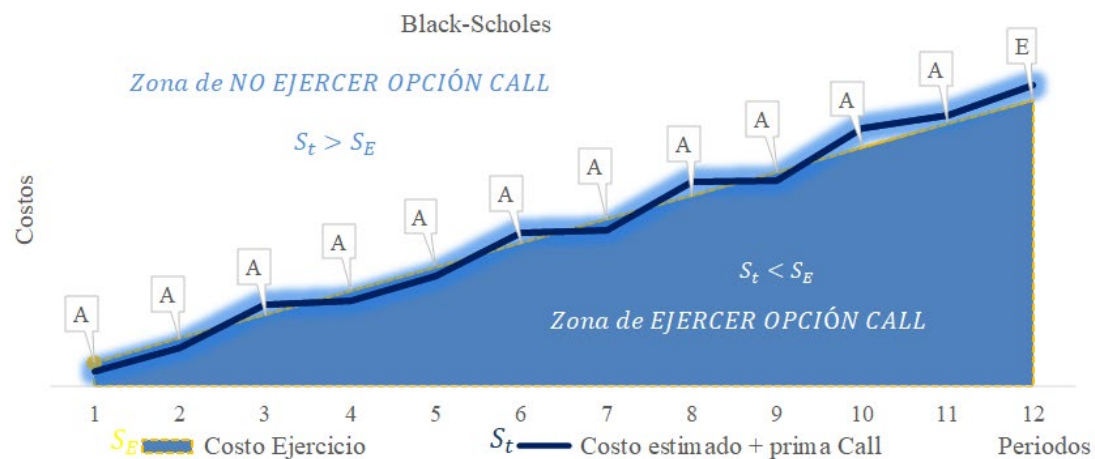
5	Modelo De Riesgo	FDP Pérdidas Mes i							
Expuestos (e)	Black-Scholes	Función de distribución	r	T	σ	d1	d2	Call	Se
29866	θ_i Mes 01	S_{GEi} Pearson5	15%	1	75%	0.56	-0.20	36,042	106,560
29773	θ_i Mes 02	S_{GEi} Pearson5	15%	1	74%	0.55	-0.19	35,211	106,560
29711	θ_i Mes 03	S_{GEi} Pearson5	15%	1	83%	0.61	-0.23	40,528	106,560
29703	θ_i Mes 04	S_{GEi} Pearson5	15%	1	73%	0.53	-0.20	34,233	106,560
29762	θ_i Mes 05	S_{GEi} Pearson5	15%	1	75%	0.57	-0.18	36,692	106,560
29737	θ_i Mes 06	S_{GEi} Pearson5	15%	1	78%	0.60	-0.18	38,830	106,560
29805	θ_i Mes 07	S_{GEi} Pearson5	15%	1	76%	0.56	-0.20	36,288	106,560
29831	θ_i Mes 08	S_{GEi} Pearson5	15%	1	76%	0.61	-0.15	39,240	106,560
29864	θ_i Mes 09	S_{GEi} Pearson5	15%	1	73%	0.56	-0.17	35,935	106,560
29932	θ_i Mes 10	S_{GEi} Pearson5	15%	1	75%	0.63	-0.12	40,283	106,560
29980	θ_i Mes 11	S_{GEi} Pearson5	15%	1	75%	0.58	-0.17	37,123	106,560
29873	θ_i Mes 12	S_{GEi} Pearson5	15%	1	78%	0.61	-0.17	39,365	106,560

Modelo de Riesgo	Criterio Opción de compra Call		Criterio Akaike Mes i		FDP Pérdidas grupo etáreo i	Prima Pura	Media St	Media St	Modelo Pérdidas Agregadas	PRIMA PURA
Black-Scholes	PP > E: No ejerce		Valor criterio		FDP Simulada	Convolución	E[S]	E[S] + Call	LDA	
θ_i Mes 01	EJERCE	0	1.26E+06	Aceptada	104,382	104,382	105,061	141,102	51646877478.6	144328.5
θ_i Mes 02	EJERCE	0	1.19E+06	Aceptada	104,909	104,909	104,828	140,040		
θ_i Mes 03	No ejerce	40,528	1.32E+06	Aceptada	108,245	108,245	107,466	147,994		
θ_i Mes 04	EJERCE	0	1.25E+06	Aceptada	104,168	104,168	103,538	137,771		
θ_i Mes 05	EJERCE	0	1.31E+06	Aceptada	106,825	106,825	106,307	142,999		
θ_i Mes 06	No ejerce	38,830	1.28E+06	Aceptada	108,552	108,552	107,724	146,554		

θ_i Mes 07	EJERCE	0	1.29E+06	Aceptada	105,901	105,901	105,170	141,458		
θ_i Mes 08	No ejerce	39,240	1.28E+06	Aceptada	107,828	107,828	109,266	148,506		
θ_i Mes 09	EJERCE	0	1.24E+06	Aceptada	106,545	106,545	105,930	141,865		
θ_i Mes 10	No ejerce	40,283	1.24E+06	Aceptada	105,285	105,285	111,254	151,536		
θ_i Mes 11	No ejerce	37,123	1.16E+06	Aceptada	106,912	106,912	106,974	144,096		
θ_i Mes 12	No ejerce	39,365	1.19E+06	Aceptada	109,164	109,164	108,633	147,998		

Fuente: Elaboración propia.

Ilustración 1. Opción call de Black-Scholes.



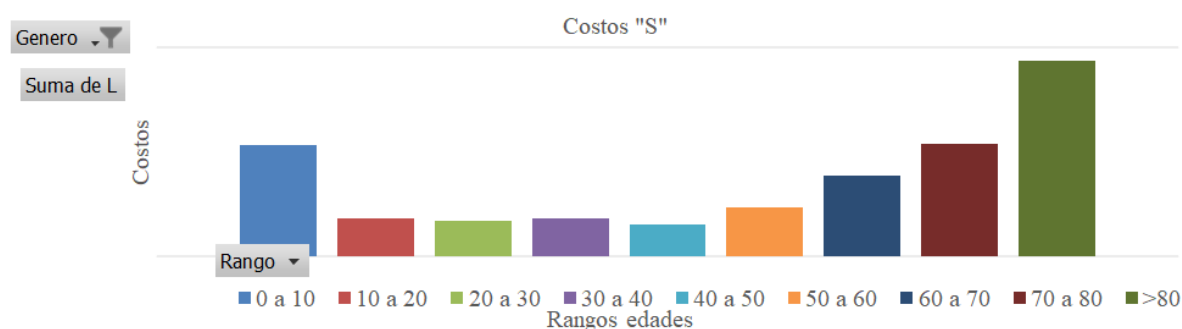
Fuente: Elaboración propia.

6. Conclusión

Los modelos aplicados para la prima de la tarifa permiten al actuario tener más campo de decisión a la hora de tarifar. En esta investigación se concluye que todos los modelos se pueden ajustar al momento de requerir estimar una tarifa, solo hace falta conocer muy bien el negocio al cual aplicar el modelo de tal forma que las variables puedan ser las apropiadas a la hora de calcular primas óptimas para cubrir con gran eficiencia las futuras pérdidas ocasionadas por los siniestros.

La Ilustración 2 fue un apoyo fundamental para analizar ciertas metodologías: mostró la necesidad de comprender el modelo de la teoría de fallas para aplicar correctamente el modelo. Mostrando con claridad lo que en el modelo era necesario, las tres etapas: de iniciación, de meseta y la etapa de finalización, formando en su conjunto la curva de la bañera y soportado desde lo estadístico con los resultados de la tarificación como se puede apreciar en la ilustración 3, con la mayor presencia en los niveles de riesgo 1, 8 y 9.

Ilustración 2. Costos por niveles de riesgo.



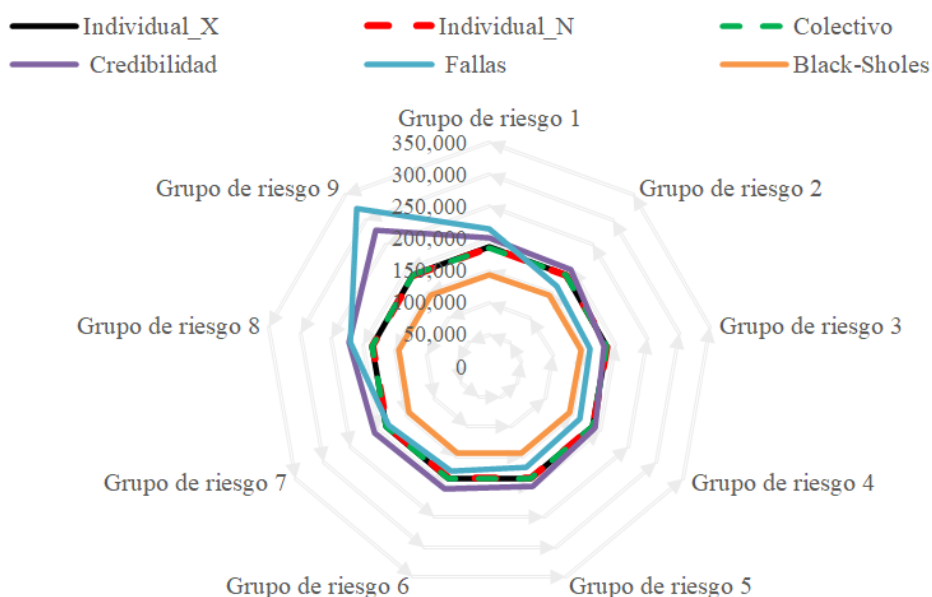
Fuente: Elaboración propia.

En la Ilustración 3, tres métodos, el método individual para las reclamaciones, el método individual para los costos y el método colectivo, se encontraron en una misma línea de prima. De tal manera que los resultados en sus distribuciones se ajustaron con primas similares. Los modelos de credibilidad y teoría de fallas presentaron unos resultados llamativos, desde el punto de vista de volatilidad, particularmente, para el grupo de riesgo 9, en este grupo se exhibe el mayor costo en las primas.

Para el modelo de Black-Scholes, las primas podrían no ser suficientes si la situación real cambiara; en tal caso, las primas, eventualmente, pueden ser superadas por las expectativas, y, a su vez, pueda fallar el principio de suficiencia. Esto comparado con el nivel de tarificación de los demás modelos.

Sin embargo, a pesar de que el modelo de Black Scholes no es el óptimo, dentro de los modelos analizados, es de rescatar como estos modelos diferentes, con respecto de los modelos utilizados en la actuaria, pueden presentar una buena aproximación a la hora de calcular tarifas y, seguramente, el modelo Black-Scholes, con un mejor ajuste, se puede incluir como una metodología para tener en cuenta.

Ilustración 3. Primas por modelos y niveles de riesgo.



Fuente: Elaboración propia.

Referencias

- Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of political economy*, 81(3), 637-654.
- Bühlmann, H. (1969). Experience rating and credibility. *ASTIN Bulletin: The Journal of the IAA*, 5(2), 157-165.
- Chen, Y., Cheung, K.C., Choi, H.M., & Yam, S. C. (2020). Evolutionary credibility risk premium. *Insurance: Mathematics and Economics*, 93, 216-229.
- Committee on Ratemaking Principles (Mayo de 1988). *Statement of Principles Regarding Property and Casualty Insurance Ratemaking*. https://www.casact.org/education/spring/2012/handouts%5CSession_5134_handout_401_0.pdf
- Cossette, H., Marceau, E., & Mtalai, I. (2019). Collective risk models with dependence. *Insurance: Mathematics and Economics*, 87, 153-168.
- Denuit, M.M., Kiriliouk, A., & Segers, J. (2015). Max-factor individual risk models with application to credit portfolios. *Insurance: Mathematics and Economics*, 62, 162-172.
- Efron, B., & Gong, G. (1983). A Leisurely Look at the Bootstrap, the Jackknife, and Cross-Validation. *The American Statistician*, 37(1), 36-48.
- Finger, R.L. (2001). Risk Classification. En *Foundations of Casualty Actuarial* (4th ed., pp. 287-342). Casualty Actuarial Society.
- Fisher, R.A. (1955). Statistical Methods and Scientific Induction. *Journal of the Royal Statistical Society*, 17(1), 69-78.

- Garrido, J., & Genest, C. (2016). Generalized linear models for dependent frequency and severity of insurance claims. *Insurance: Mathematics and Economics*, 70, 205-215.
- Gulumsen, M., Tonkin, R.S., & Johannes, D.J. (2002). Competition in the general insurance industry. *Zeitschrift für die gesamte Versicherungswissenschaft*, 91(3), 453-481.
- Hua, L. (2015). Tail negative dependence and its applications for aggregate loss. *Mathematics and Economics*, 61, 135-145.
- Jeong, H., & Valdez, E.A. (2020). Predictive compound risk models with dependence. *Insurance: Mathematics and Economics*, 94, 182-195.
- Jiang, R. (2013). A new bathtub curve model with a finite support. *Reliability Engineering & System Safety*, 119, 44-51.
- Klugman, S.A., Panjer, H.H., & Willmot, G. E. (2008). *Loss Models. From Data to Decisions* (3th ed.). Hoboken, New Jersey: John Wiley & Sons, Inc.
- Klugman, S.A., Panjer, H. H., & Willmot, G.E. (2019). *Loss Models: From Data to Decisions* (5th ed.). New Jersey: John Wiley and Sons, Inc.
- Lamothe, P., & Pérez, M. (2005). *Opciones financieras y productos estructurados* (3th ed.). Madrid: McGraw-Hill Interamericana.
- Lee, G.Y., & Shi, P. (2019). A dependent frequency-severity approach to modeling longitudinal insurance claims. *Insurance: Mathematics and Economics*, 87, 115-129.
- Lenhard, J. (2006). Models and Statistical Inference: The Controversy between Fisher and Neyman-Pearson. *The British Journal for the Philosophy of Science*, 57(1), 69-91.
- Marín, J.M., & Rubio, G. (2001). *Economía financiera*. Antoni Bosch editor.
- Martel, M., Hernández, A., & Vázquez, F.J. (2012). On the independence between risk profiles in the compound collective risk actuarial model. *Mathematics and Computers in Simulation*, 82(8), 1419-431.
- Migon, H.S., & Moura, F.A. (2005). Hierarchical Bayesian collective risk model: an application to health insurance. *Insurance: Mathematics and Economics*, 36(2), 119-135.
- Miller, R. B., & Hickman, J.C. (1975). Teoría de la credibilidad del seguro y estimación bayesiana. *Teoría y aplicaciones de la credibilidad*, 249-270.
- Minsalud. (2020). *Estudio de suficiencia y de los mecanismos de ajuste del riesgo para el cálculo de la Unidad de Pago por Capitación, recursos para garantizar la financiación de tecnologías en salud y servicios en los regímenes Contributivo y Subsidiado*. Colombia: Ministerio de la Salud.
- Monterrey, P. (2012). $P < 0,05$, ¿Criterio mágico para resolver cualquier problema o leyenda urbana? *Universidad Javeriana*, 17(2), 203-215.
- Moreno, M.T., & Ramos, L. (2003). *Aplicación de Modelos de Credibilidad para el Cálculo de Primas en el Seguro de Automóviles*. Comisión Nacional de Seguros y Fianzas de los Estados Unidos Mexicanos.

- Mudholkar, G.S., Asubonteng, K.O., & Hutson, A.D. (2009). Transformation of the bathtub failure rate data in reliability for using Weibull-model analysis. *Statistical Methodology*, 6(6), 622-633.
- Norberg, R. (1979). The credibility approach to experience rating. *Scandinavian Actuarial Journal*, 4, 181-221.
- OCDE. (2016). Principios de Gobierno Corporativo de la OCDE y del G20. (p. 68). Paris: OCDE. <http://dx.doi.org/10.1787/9789264259171-es>
- Parodi, P. (2015). *Pricing in General Insurance* (Vol. 10). New York: Chapman and Hall.
- Roesch, W. J. (2012). Using a new bathtub curve to correlate quality and reliability. *Microelectronics Reliability*, 52(12), 2864-2869.
- Roos, B. (2007). On variational bounds in the compound Poisson approximation of the individual risk model. *Insurance: Mathematics and Economics*, 40(3), 403-414.
- Shreve, S.E. (2004). *Stochastic calculus for finance II: Continuous-time models* (Vol. 11). New York: Springer.
- Schinzinger, E., Denuit, M.M., & Christiansen, M.C. (2016). A multivariate evolutionary credibility model for mortality improvement rates. *Insurance: Mathematics and Economics*, 69, 70-81.
- Tsai, C.C.-L., & Wu, A.D. (2020). Incorporating hierarchical credibility theory into modelling of multi-country mortality rates. *Insurance: Mathematics and Economics*, 91, 37-54.
- Wen, L., Wu, X., & Zhou, X. (2009). The credibility premiums for models with dependence induced by common effects. *Insurance: Mathematics and Economics*, 44(1), 19-25.
- Werner, G., & Modlin, C. (2019). *Basic Ratemaking*. Casualty Actuarial Society.
- Xie, M., Tang, Y., & Goh, T. N. (2002). A modified Weibull extension with bathtub-shaped failure rate function. *Reliability Engineering & System Safety*, 76(3), 279-285.
- Yang, J., Zhou, S., & Zhang, Z. (2005). The compound Poisson random variable's approximation to the individual risk model. *Insurance: Mathematics and Economics*, 36(1), 57-77.
- Yeo, K.L., & Valdez, E.A. (2006). Claim dependence with common effects in credibility models. *Insurance: Mathematics and Economics*, 38(2), 609-629.
- Zhang, T., Xie, M., Tang, L.C., & NG, S.H. (2005). Reliability and modeling of systems integrated with firm ware and hard ware. *International Journal of Reliability, Quality and Safety Engineering*, 12(3), 227-339.