



TRABAJO DE FIN DE MÁSTER

MÁSTER EN CIENCIAS ACTUARIALES Y FINANCIERAS

**MÉTODOS AVANZADOS PARA TARIFICACIÓN
DE SEGUROS NO VIDA: APLICACIÓN DEL
ALGORITMO RECURSIVO DE PANJER PARA
LA CONSTRUCCIÓN DE UN MODELO
ESTADÍSTICO DE PÉRDIDAS AGREGADAS**

Autor

Eugenio Margallo Pereira

Tutores

D. Jesús Ramón Simón del Potro

D. José Miguel Rodríguez-Pardo del Castillo

Curso Académico 2017/18



Esta tesis es propiedad del autor. No está permitida la reproducción total o parcial de este documento sin mencionar su fuente. El contenido de este documento es de exclusiva responsabilidad del autor, quien declara que no se ha incurrido en plagio y que la totalidad de referencias a otros autores han sido expresadas en el texto.

En caso de obtener un calificación igual o superior a 8.0 Notable, autorizo la publicación de este trabajo en el centro de Documentación de la Fundación Mapfre.

- Sí, autorizo a su publicación.
- No, desestimo su publicación.

Firmado: Eugenio Margallo Pereira



Agradecimientos

Quisiera aprovechar la oportunidad de redactar el presente Trabajo de Fin de Máster para agradecer a mi familia el gran esfuerzo e incontables sacrificios que han realizado durante tantos años por darme la oportunidad de estudiar y apoyarme incondicionalmente durante todo este tiempo. Gracias por transmitirme unos grandes valores y exigirme siempre para superarme y no conformarme con el mínimo. En especial, eternamente gracias a mi abuela Francisca por todo lo que me enseñó desde el punto de vista más sencillo y humilde posible.

También a mis amigos, por su apoyo y ánimos en las etapas difíciles y por los grandes momentos que hemos compartido cuando nuestros compromisos nos lo han permitido. Marta, gracias por todo, por tu inmenso cariño y apoyo en estos años.

No puedo olvidarme en esta ocasión de todos los profesores que he tenido desde que empecé a estudiar. Todos me habéis aportado algo por pequeño que fuera, y ayudado a ir cumpliendo las metas que me he ido proponiendo a lo largo de estos años. Especialmente, agradecer al profesor José María Serrano su entrega y perseverancia en las clases de Estadística II de las que tanto aprendí.

Por último, gracias a todos los compañeros de trabajo, desde el Hotel Santo Domingo a PricewaterhouseCoopers pasando por Trevo y AESMIDE; de todos vosotros he aprendido mucho.



Resumen

El proceso de tarificación en seguros No Vida está sujeto al conocimiento del comportamiento de los asegurados, atendiendo a la experiencia pasada. Esto implica estudiar variables relacionadas con el comportamiento de individuos de una cartera y las características que presenta cada subconjunto de ésta.

Para ello, se recurre a la modelización estadística que permita definir una distribución de probabilidad que, bajo determinados parámetros, pudiera explicar de la mejor manera el comportamiento, tanto de frecuencia de siniestros como de los costes que implican.

Con el fin de analizar el comportamiento de cada subgrupo, se recurre al algoritmo recursivo de Panjer para construir, con las distribuciones de probabilidad que mejor representen el comportamiento, una nueva distribución y utilizarla para definir la prima. Esta distribución puede ser utilizada también para evaluar la adecuación de la tarifa y los riesgos del producto.

Palabras clave: Distribución de probabilidad, parámetro, frecuencia, severidad, clase (a,b,0), clase (a,b,1), estadísticos descriptivos, algoritmo de Panjer, prima, severidad simulada.

Abstract

Pricing process of Nonlife insurance is based on the knowledge of the behaviour of the assured by using past experience. This implies studying variables linked with the behaviour of individuals who comprise a portfolio and the characteristics presented by each subset of it.

To this end, it is necessary the use of statistical modelling which allows to define a probability distribution that, under specific parameters, could explain the behaviour of the assured in a more representative way, both frequencies and severities involved.

For the purpose of analysing the behaviour of each cluster, we apply Panjer recursion algorithm using the distributions that better fit the behaviour of the clients, in order to create a new distribution. This new distribution is then used to calculate the premium. This new distribution can be also used to assess the appropriateness of the premium and risk of the product.

Key Words: Statistical distribution, parameter, frequency, severity, (a,b,0) class, (a,b,1) class, descriptive statistics, Panjer's algorithm, premium, simulated severity.



Índice

1. Presentación del Trabajo de Fin de Máster	7
Introducción	7
Objetivos	8
Descripción del trabajo	9
Resultados alcanzados	10
Motivaciones personales	10
2. Seguros No Vida: Presentación y descripción	11
3. Análisis de los datos de asegurados	12
3.1. Estudio de frecuencias de siniestros	13
3.1.1. Distribuciones discretas de la clase (a, b, o).....	14
Distribución Binomial.....	15
Distribución Binomial Negativa	17
Distribución Geométrica.....	19
Distribución Poisson.....	20
3.1.2. Distribuciones discretas de la clase (a, b, 1)	22
3.2. Estudio de severidades de siniestros.....	24
Distribución Gamma.....	25
Distribución Exponencial	26
Distribución Pareto Generalizada.....	28
Distribución Inversa Gaussiana.....	30
Distribución Logistic.....	32
Distribución LogLogistic.....	33
Distribución LogNormal	35
Distribución Normal	37
Distribución Weibull.....	38
4. Elección de la distribución de probabilidad de frecuencias y severidades	41
4.1. Estimación de parámetros de la distribución por Máxima Verosimilitud	41
4.1.1. Método de Máxima Verosimilitud para casos discretos	42
4.1.2. Método de Máxima Verosimilitud para casos continuos	43
4.2. Realización del Test Chi-Cuadrado: Bondad del ajuste.....	43
4.3. Comprobación del p-Valor	45
5. Distribución de reclamaciones totales mediante el método de Panjer	46
5.1. Discretización de las distribuciones continuas: Método <i>Rounding</i>	46



5.2.	Aplicación del algoritmo recursivo de Panjer	47
6.	Tarificación de seguros No Vida	51
6.1.	<i>Value-at-Risk</i> : medida de riesgo para la tarificación de seguros No Vida	52
6.2.	Principio general para la tarificación de seguros y sus extensiones	53
7.	Simulación de severidad: Adecuación de las primas cobradas	55
8.	Resultados alcanzados: Ejemplo ilustrativo	57
8.1.	Inputs del aplicativo	57
8.2.	Resultados del análisis y tarificación	59
8.2.1.	Resultados de ajustes a distribución de probabilidad teóricas	59
8.2.2.	Obtención de la distribución de pérdidas agregas por el algoritmo de Panjer.....	60
8.2.3.	Tarificación de seguros para cada clúster	63
8.2.4.	Resumen de la cartera del producto asegurador	64
9.	Conclusiones finales	65
	Bibliografía	67

Índice de Tablas

Tabla 1.	Resumen de los parámetros “a” y “b” y cálculo de $f_x(0)$ de las distribuciones discretas de la clase (a,b,0).....	15
Tabla 2.	Resumen de los parámetros “a” y “b” y cálculo de $f_x(0)$ de las distribuciones discretas de la clase (a,b,1).....	24
Tabla 3.	Cálculo del valor de $f_s(0)$ según la distribución de probabilidad de las frecuencias. ...	49

Índice de Gráficos

Gráfico 1.	Función de probabilidad de la distribución Binomial (n,p).....	16
Gráfico 2.	Función de distribución de la distribución Binomial (n,p).....	17
Gráfico 3.	Función de probabilidad de la distribución Binomial Negativa (r,p).....	18
Gráfico 4.	Función de distribución de la distribución Binomial Negativa (r,p).....	18
Gráfico 5.	Función de probabilidad de la distribución Geométrica (p).....	20
Gráfico 6.	Función de distribución de la distribución Geométrica (p).....	20
Gráfico 7.	Función de probabilidad de la distribución Poisson (λ).....	21
Gráfico 8.	Función de distribución de la distribución Poisson (λ).....	22
Gráfico 9.	Función de probabilidad de la distribución Gamma (α, β).....	26
Gráfico 10.	Función de distribución de la distribución Gamma (α, β).....	26
Gráfico 11.	Función de probabilidad de la distribución Exponencial (μ).....	27



Gráfico 12. Función de distribución de la distribución Exponencial (μ).....	28
Gráfico 13. Función de probabilidad de la distribución Pareto Generalizada (k, σ, θ).	29
Gráfico 14. Función de distribución de la distribución Pareto Generalizada (k, σ, θ).	30
Gráfico 15. Función de probabilidad de la distribución Inversa Gaussiana (μ, λ).....	31
Gráfico 16. Función de distribución de la distribución Inversa Gaussiana (μ, λ).....	31
Gráfico 17. Función de probabilidad de la distribución Logistic (μ, σ).	33
Gráfico 18. Función de distribución de la distribución Logistic (μ, σ).	33
Gráfico 19. Función de probabilidad de la distribución LogLogistic (μ, σ).	34
Gráfico 20. Función de distribución de la distribución LogLogistic (μ, σ).....	35
Gráfico 21. Función de probabilidad de la distribución LogNormal (μ, σ).	36
Gráfico 22. Función de distribución de la distribución LogNormal (μ, σ).	36
Gráfico 23. Función de probabilidad de la distribución Normal (μ, σ).	38
Gráfico 24. Función de distribución de la distribución Normal (μ, σ).	38
Gráfico 25. Función de probabilidad de la distribución Weibulll (A, B).	39
Gráfico 26. Función de distribución de la distribución Weibull (A, B).	40

Índice de Imágenes

Imagen 1. Introducción de inputs para análisis de clústeres y tarificación de seguros.	58
Imagen 2. Ajuste de distribuciones de frecuencias y severidades de cada clúster.....	59
Imagen 3. Ajuste de distribuciones a valores observados de cada clúster.....	60
Imagen 4. Resultados de cálculo de la distribución de pérdidas agregadas mediante método de Panjer.....	61
Imagen 5. Distribución de pérdidas agregadas por clúster (izquierda) y del producto asegurador (derecha).	62
Imagen 6. Resultados de la tarificación de cada clúster.	63
Imagen 7. Resumen de la tarificación global del riesgo	64



1. Presentación del Trabajo de Fin de Máster

Introducción

El negocio de las entidades aseguradoras está basado en la asunción de riesgos a cambio de la percepción de una prima, la cual queda establecida al firmar el contrato entre tomador del seguro y asegurador, previo estudio por parte de la entidad aseguradora del riesgo a cubrir. El contrato de seguro supone un acuerdo que implica el intercambio de capital en forma de unidades monetarias entre el asegurado y la compañía aseguradora. La parte tomadora transfiere un riesgo de naturaleza aleatoria a cambio de un precio conocido, mientras que el asegurador, por su parte, recibe una cuantía con la que deberá asumir el posible acaecimiento de las contingencias que se recojan en el contrato y cuyo coste estará sujeto a aleatoriedad.

Dentro de los negocios de las aseguradoras se encuentran dos grandes ramos: Vida y No Vida. Los seguros vinculados al ramo de Vida son aquellos en los que las cuantías recogidas en el contrato están sujetos a la supervivencia, fallecimiento o invalidez del asegurado. Estos eventos pueden estudiarse a través de factores biométricos como la edad, lo cual ayuda a aproximar una probabilidad de muerte, supervivencia o invalidez según la experiencia de otros grupos recogidos en tablas biométricas.

Por otro lado, los seguros que se enmarcan dentro del ramo de No Vida son aquellos en los que la ocurrencia de un determinado acontecimiento no puede modelizarse atendiendo a esta base biométrica. Dentro del ramo No Vida se encuentran los seguros de autos, seguros del hogar, seguros de robo, seguros de viajes, etc.

Para poder comprender en profundidad los riesgos inherentes en los contratos que se suscriben en las múltiples líneas de negocio que abarca el ramo de No Vida, se estudia la experiencia de los individuos a lo largo del tiempo. Esto se realiza con el fin de detectar patrones de comportamiento de los individuos que conforman la cartera del producto asegurador, definiendo habitualmente subgrupos homogéneos o clústeres que compartan determinadas características. Estos subgrupos suelen constituirse siguiendo distintos criterios de segmentación que las entidades aseguradoras utilizan para identificar a sus asegurados (según su edad, nivel de estudios, zonas geográficas de residencia, ingresos, etc.).

Gracias a la constitución de clústeres con características comunes y al estudio del comportamiento pasado de los asegurados, es posible desarrollar una tarifa de seguro ajustada al



riesgo que éstos presentan, ya que esos datos ofrecen información de los siniestros y los costes que el asegurado ha presentado hasta ese momento, revelando así patrones de comportamiento.

Por otro lado, para poder garantizar que la compañía aseguradora puede hacer frente a los compromisos adquiridos al firmar un contrato, se deberán constituir unas reservas que sean suficientes para responder al objeto del seguro.

En cualquier caso, hay que considerar que la manera de entender el negocio asegurador se basa en reducir los riesgos asumidos mediante agregación, es decir, cuantas más pólizas se incluyan en la cartera de la compañía, los riesgos decrecerán. Esto es debido a que es poco probable que todos los asegurados reclamen a la vez, por lo que, al incluir un mayor número de pólizas que la aseguradora suscribe, la variabilidad que presente será menor y se podrá hacer frente a las cuantías reclamadas.

Objetivos

Una vez realizada esta breve introducción al negocio asegurador, el presente Trabajo de Fin de Máster (de aquí en adelante, TFM), del Máster en Ciencias Actuariales y Financieras de la Universidad Carlos III de Madrid, tiene como objetivo el desarrollo de un aplicativo del perfil de seguros No Vida desarrollado con el software MATLAB. Este aplicativo tiene como meta principal la tarificación de seguros del ramo No Vida, por lo que se debe cumplir los siguientes objetivos intermedios para alcanzarla:

1. Análisis estadístico de los datos presentados por cada uno de los grupos homogéneos.
2. Selección de las distribuciones de probabilidad que mejor representen el comportamiento de los individuos de cada clúster.
3. Construcción mediante el método de Panjer de la distribución de pérdidas agregadas.
4. Simulación de severidades y análisis del *Value-at-Risk (VaR)* de cada grupo homogéneo.
5. Desarrollo de tarifa en base a la distribución de pérdidas agregadas de Panjer.
6. Evaluación y análisis de la consistencia de las primas cobradas a cada grupo homogéneo frente a las severidades simuladas.
7. Evaluación y análisis global de la tarificación y viabilidad económica del producto asegurador analizado.

El disponer de datos relacionados con la frecuencia de siniestros y severidad de cada clúster permite establecer una tarifa que ofrezca cobertura a un determinado riesgo siendo el precio lo más adecuado posible a cada grupo.



Adicionalmente, se determina un resultado, mediante simulación siguiendo los datos de las severidades agregadas, de las posibles reclamaciones de los clústeres y, por extensión, de la cartera. Este resultado puede aproximar la severidad que la entidad experimentará durante la vigencia del contrato. El departamento actuarial dará razonabilidad a dicho resultado para dotar las reservas adecuadas. También, el departamento de gestión de riesgos puede utilizar los resultados de la severidad simulada ofrecidos por el aplicativo para evaluar la adecuación de la comercialización del producto y apoyar a la estrategia de la entidad, entre otras aplicaciones.

Descripción del trabajo

Para poder desarrollar este aplicativo de tarificación de seguros No Vida y alcanzar los objetivos anteriores, es necesario elaborar una serie de estudios de los datos que se utilizan como base. En el presente TFM se utilizarán datos de frecuencias de siniestros y severidades asociadas a los individuos que pertenecen al clúster de un periodo de tiempo específico. Seguidamente, se procederá a ajustar el comportamiento de ambas variables a distintas distribuciones de probabilidad, utilizando distribuciones discretas para las frecuencias y distribuciones continuas para las severidades.

Una vez identificada la distribución de probabilidad que mejor se ajusta tanto para las frecuencias como para las severidades de cada clúster, se utilizará el método recursivo de Panjer para obtener la distribución de reclamaciones totales. Mediante este algoritmo, se calculará la distribución de las pérdidas agregadas exacta y será la base para posteriormente efectuar la tarificación de cada clúster.

Una vez obtenida la distribución de pérdidas agregadas de cada clúster, se estudian los estadísticos media y desviación típica de cada uno. Utilizando la distribución de pérdidas agregadas obtenida por el algoritmo, se calculará el *Value-at-Risk* de cada clúster y será, junto con la media, los pilares para el proceso de *pricing* de seguros del aplicativo.

Finalmente, como complemento a los cálculos de tarificación efectuados por el aplicativo, se calcula la función de distribución de pérdidas agregadas de cada clúster y se procede a realizar simulaciones de escenarios con números aleatorios con el fin de conseguir un valor para evaluar lo adecuada que es la prima cobrada y qué cantidad podría ser reclamada en un determinado escenario.

Para poder desarrollar el presente TFM se utilizará material didáctico presentado durante el Máster en las asignaturas de Estadística Actuarial y Tarificación de Seguros No Vida. Se recurrirá



a literatura especializada recomendada por los tutores y distintos recursos procedentes de Internet como soporte al código desarrollado en MATLAB y a conceptos estadístico-actuariales.

Resultados alcanzados

A lo largo del presente documento, se explicará más en detalle la metodología aplicada, así como la teoría y matemática actuarial que soporta los cálculos. Además, se acompaña el documento con capturas de los resultados ofrecidos por el aplicativo, explicando éstos en cada paso de los anteriormente citados para facilitar la comprensión.

Motivaciones personales

El motivo por el cual me decanto en enfocar mi Trabajo de Fin de Máster a esta temática es por mi especial interés en el área de No Vida. Una vez superado el curso y habiendo adquirido nuevos conocimientos en las asignaturas de Estadística Actuarial y Tarificación de Seguros No Vida, consideré el desarrollo de un aplicativo en MATLAB como un reto el cual me ha permitido asentar conceptos estadísticos, actuariales y de programación.

La elección de este tema me ha ayudado a comprender cómo es y la importancia que cobra el proceso de tarificación de seguros. En concreto, ahondar en las principales características de las distribuciones de probabilidad que suelen utilizarse en el ámbito actuarial a la hora de realizar el *pricing* de un producto me ha ayudado a entender la modelización de las primas para los clústeres con el enfoque propuesto por Panjer.

Por otro lado, el construir el software en una aplicación como MATLAB ha permitido recurrir a un amplio abanico de recursos en línea para conseguir un funcionamiento del aplicativo óptimo y poder desarrollar posteriormente una interfaz gráfica de utilización y comprensión sencilla para el usuario sobre un ejecutable que no requiera tener instalado el software en el equipo.



2. Seguros No Vida: Presentación y descripción

Los seguros de No Vida son aquellos que protegen al tomador del seguro o asegurado frente a las posibles pérdidas financieras que surgen como consecuencia de la ocurrencia de eventos de naturaleza aleatoria. Estos seguros, al contrario de los seguros del ramo Vida, no están sujetos a la supervivencia, fallecimiento o invalidez del contratante o asegurado.

Algunos ejemplos de seguros de No Vida serán los de autos (daños y accidentes de coches), hogar (daños sufridos en un domicilio por diversas causas: incendios, rotura de cristales, humedades...), robos, catastróficos (incendios, terremotos, inundaciones...) o viaje, entre muchos otros. En ellos, a priori, no habría una forma de vincular la posible ocurrencia del evento de manera directa a las características biométricas del tomador del seguro, pero sí es posible estudiar patrones de comportamiento de los clientes que componen la cartera de la entidad. De este modo, es posible desagregar la cartera de un producto en cuestión en subgrupos que contengan una población lo suficientemente extensa como para desarrollar estudios sobre ella.

En el caso de los seguros de No Vida, el proceso de tarificación puede adquirir cierta complejidad al estar sometido a un análisis estadístico que puede suponer cierto nivel de exigencia técnica en los cálculos. Como se ha explicado anteriormente, la tarificación no puede realizarse a priori de manera directa atendiendo a características biométricas. Sin embargo, puede realizarse mediante el estudio del comportamiento de las variables clave para la tarificación, identificándolas y modelizándolas en cada subgrupo.

En el presente TFM se tomarán como variables la frecuencia siniestral, que representa el número de eventos que el tomador ha presentado, y la severidad de los siniestros, es decir, el coste que suponen los siniestros. Además, para poder identificar los clústeres, se añadirá un número identificativo a cada asegurado de la cartera que corresponderá al clúster en el cual la entidad aseguradora ha clasificado a ese cliente según las características que considere oportunas.



3. Análisis de los datos de asegurados

Las compañías aseguradoras que operan en el ramo No Vida utilizan un gran volumen de variables para estudiar y agrupar a sus asegurados en clústeres. De esta manera, la compañía tiene una población amplia de clientes identificada y que presentan rasgos de comportamiento muy similares, lo cual permite fijar una tarifa lo más ajustada y adecuada posible para esa población concreta.

Si bien las variables que se utilizan para caracterizar a los clientes pueden obedecer a distintos criterios, la clusterización es un campo de estudio de la Estadística Multivariante y cada empresa puede recurrir a distintas metodologías para efectuarla, atendiendo además a criterios previos de segmentación de carteras de asegurados. Entre los distintos métodos para el *clustering* están los jerárquicos (con asociaciones basadas en distancias entre los valores: simples, promedias y completas), el criterio de Ward (atendiendo a una partición en k grupos) o métodos no jerárquicos (aplicando el algoritmo de las *k-medias*). En consecuencia, el proceso de clusterización es un proceso propio de cada entidad y no será aplicado en el presente TFM. Por ello, se opta por importar datos de asegurados añadiendo una variable adicional que identifique a cada asegurado de la base de datos con un clúster, que será un número (1, 2, 3...). De esta forma se asume que la base ha sido analizada previamente por la entidad según su propio criterio.

Una vez identificado el clúster, las dos variables que se analizarán serán la frecuencia de siniestros y la severidad. La primera presenta las características de una variable aleatoria discreta, mientras que la segunda se estudia como una variable aleatoria continua.

Se entiende por variable aleatoria discreta aquellas que sólo pueden tomar un número finito de valores. Por ejemplo, los siniestros que se presentan en una compañía aseguradora podrían ser 0, 1, 2, 3... pero nunca podrá darse un siniestro entre 1 y 2, por lo que se entiende esta variable como discreta y se modelizará con distribuciones de probabilidad acordes.

Por otra parte, la severidad representa el coste del siniestro presentado por el asegurado y por el cual, en virtud del contrato, la entidad debe responder por él si se cumplen las condiciones recogidas en la póliza. En este caso, se entiende la severidad como una variable continua, puesto que entre dos valores puede tomar un rango de valores intermedio. Supongamos que se presentan severidades por valor de 1.000 unidades monetarias y otra por 2.000 unidades monetarias. En este caso, podría presentarse cualquier cuantía intermedia entre 1.000 y 2.000 (1.200,15, 1732,98, 1031,43...). Por ese motivo, se asociará la severidad con distribuciones de probabilidad continuas.



A continuación, como paso previo a la explicación de las distribuciones de probabilidad, se realizará una breve introducción el algoritmo de Panjer para el cálculo de la distribución de pérdidas agregadas, el cual será estudiado con mayor detalle en el epígrafe 5, con el objetivo de contextualizar las explicaciones sucesivas.

La manera convencional de llevar a cabo el cálculo de la distribución de pérdidas agregadas es mediante las convoluciones. Recurriendo a este método de cálculo, se obtendría la distribución de pérdidas exacta. Sin embargo, la aplicación de la convolución requiere de un gran tiempo operacional debido al gran número de cálculos que implica esta metodología.

Para salvar esta condición, Panjer propuso un algoritmo el cual, si se conoce cuál es la distribución de probabilidad de la frecuencia de los siniestros y ésta es una Binomial, Binomial Negativa, Geométrica o Poisson, es posible aplicar el método desarrollado por Panjer para el cálculo de la distribución de pérdidas agregadas. De este modo, se obtiene la distribución de pérdidas agregadas con un grado de precisión que ofrece resultados idénticos al de la convolución y ganando mayor eficiencia en el cálculo. Además, con este método se abarcan las distribuciones más importantes y representativas dentro del mundo actuarial.

Las cuatro distribuciones discretas citadas anteriormente tienen como particularidad que pertenecen a la clase $(a,b,0)$ y también a la clase $(a,b,1)$, Esto significa que puede aplicarse una fórmula general para el cálculo de las probabilidades que toma la distribución de probabilidad en distintos puntos de la función.

A continuación, se procede a explicar con mayor detalle las distribuciones discretas y las características de las clases $(a,b,0)$ y $(a,b,1)$.

3.1. Estudio de frecuencias de siniestros

Se entiende como frecuencia de siniestros el número de siniestros que el asegurado reporta a la compañía durante un periodo de tiempo y a los que ésta debería hacer frente debido al contrato establecido entre ambas partes. Si bien la frecuencia de siniestros no permite deducir cuál será la cuantía final a la que la compañía aseguradora deberá hacer frente, sí es una de las variables clave en la obtención de una distribución de pérdidas agregadas.

Esta variable presenta las características de una variable aleatoria discreta, por lo que se ha procedido a ajustar las frecuencias siniestrales a las distribuciones más comunes e importantes en la literatura y práctica actuarial: Binomial, Binomial Negativa, Geométrica y Poisson.



3.1.1. Distribuciones discretas de la clase (a, b, o)

Las distribuciones Binomial, Binomial Negativa, Geométrica y Poisson mencionadas son distribuciones de probabilidad discreta que pertenecen a la clase (a,b,o). El hecho de pertenecer a esta clase permitirá aplicar, como brevemente se ha introducido al inicio del presente epígrafe, el algoritmo recursivo de Panjer para la obtención de la distribución de pérdidas agregadas.

Se habla de distribuciones de la clase (a,b,o) cuando dichas distribuciones cumplen la siguiente función recursiva:

$$f_x(x) = \left(a + \frac{b}{x}\right) f_x(x-1); \text{ siendo } x = 1, 2, 3, 4 \dots$$

[Fórmula 1]

Los parámetros “a” y “b” son calculados según la distribución de la cual quiera obtenerse la probabilidad. De igual manera, será necesario conocer el valor que toma $f_x(o)$. En la *Tabla 1* que figura más adelante se puede observar la fórmula de cálculo de “a”, “b” y $f_x(o)$ para cada distribución de probabilidad de la clase (a,b,o).

Puesto que el fin último de las distribuciones utilizadas es poder aplicarlas en la fórmula recursiva de Panjer, se recurre a la reparametrización de los parámetros “a” y “b” para los casos concretos de la distribución Geométrica y de la distribución Binomial Negativa.

De este modo, se utilizará la *Fórmula 2* para obtener un nuevo elemento β que permitirá realizar un adecuado casamiento de las distribuciones discretas vinculadas a la frecuencia de siniestros y la fórmula recursiva de Panjer:

$$\beta = \frac{(1-p)}{p}$$

[Fórmula 2]

A continuación, se resume en la *Tabla 1* cuál es la correspondencia y forma de cálculo de los parámetros “a” y “b” y el valor de $f_x(o)$ según la distribución discreta, incluyendo el ajuste anteriormente mencionado para la distribución Geométrica y Binomial Negativa.



Tabla 1. Resumen de los parámetros “a” y “b” y cálculo de $f_x(o)$ de las distribuciones discretas de la clase (a,b,o) .

Distribuciones	Parámetro “a”	Parámetro “b”	$f_x(o)$
Binomial	$\frac{-p}{(1-p)}$	$\frac{(n+1)p}{(1-p)}$	$(1-p)^n$
Binomial Negativa	$\frac{\beta}{1+\beta}$	$(r-1)\frac{\beta}{1+\beta}$	$\frac{1}{(1+\beta)^r}$
Geométrica	$\frac{\beta}{1+\beta}$	o	$\frac{1}{1+\beta}$
Poisson	o	λ	$e^{-\lambda}$

Fuente: Elaboración propia a partir de la información obtenida del libro “Loss Models: From data to decisions” (2012).

Seguidamente se exponen las distribuciones de probabilidad discretas que han sido mencionadas anteriormente, explicando sus principales aplicaciones y características.

Distribución Binomial

La distribución Binomial es una distribución discreta biparamétrica que utiliza los parámetros n y p para su modelización. Esta distribución se utiliza para estudiar experimentos en los que realizando un número determinado de intentos se quiere obtener la probabilidad de conseguir un número específico de casos favorables, siendo la probabilidad de éxito constante en cada intento.

Se define el parámetro n como el número de intentos que se realizan de un experimento y el parámetro p como la probabilidad de alcanzar el resultado deseado. Por tanto, el parámetro n se definirá como un número entero mayor que 0 y el parámetro p como un valor comprendido entre 0 y 1. La distribución Binomial es una extensión de la distribución de Bernoulli, siendo el experimento realizado una única vez en este último caso.

Para obtener la función de probabilidad y la función de distribución de la distribución Binomial, se recurre a las siguientes fórmulas:

$$f_x(x) = \frac{n!}{x!(n-x)!} p^x(1-p)^{n-x}; \text{ siendo } x = 0, 1, 2, 3 \dots n$$

[Fórmula 3]

$$F_x(x) = \sum_{i=0}^x \frac{n!}{i!(n-i)!} p^i(1-p)^{n-i}$$

[Fórmula 4]

Los estadísticos de la distribución se calculan de la manera que sigue a continuación:

$$E(X) = np$$

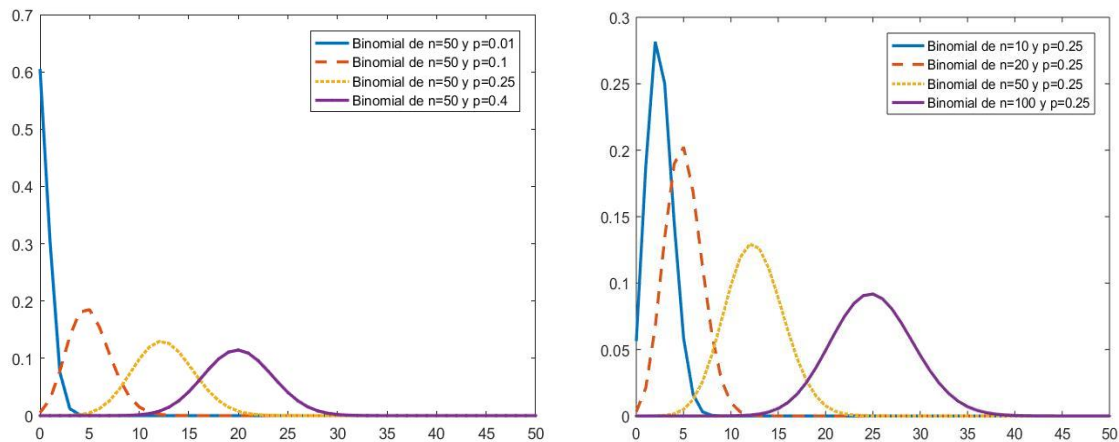
[Fórmula 5]

$$Var(X) = np(1-p)$$

[Fórmula 6]

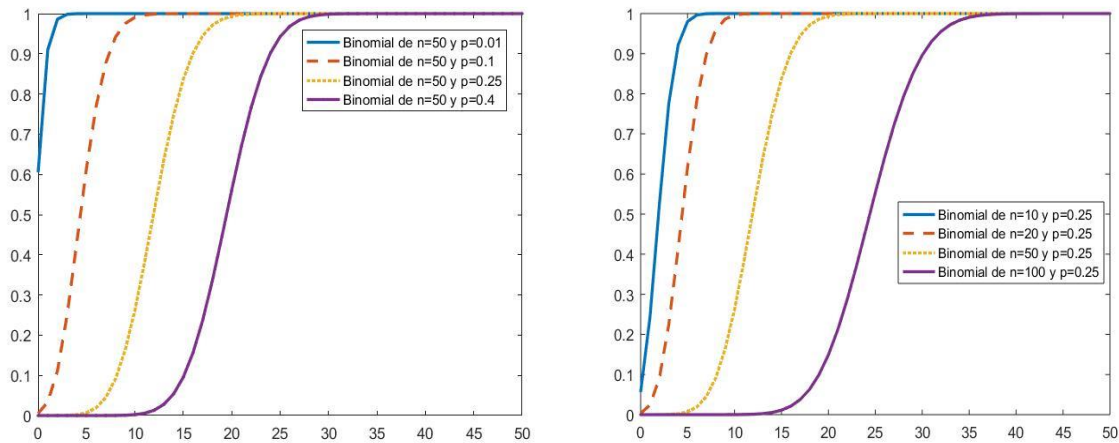
A continuación, se muestran algunos ejemplos gráficos del comportamiento de la distribución Binomial alternando los valores de los parámetros n y p .

Gráfico 1. Función de probabilidad de la distribución Binomial (n,p).



Fuente: Elaboración propia.

Gráfico 2. Función de distribución de la distribución Binomial (n,p).



Fuente: Elaboración propia.

Distribución Binomial Negativa

La distribución Binomial Negativa es una distribución biparamétrica de parámetros r y p , siendo un modelo adecuado para procesos en los que se repite un experimento en sucesivas ocasiones y el objetivo se centra en estudiar la probabilidad de obtener un número específico de fracasos hasta que se alcance el r -ésimo éxito.

El parámetro r es un número entero positivo que representa el número de éxitos del experimento y el parámetro p como la probabilidad asociada al experimento realizado, por lo que se definirá entre 0 y 1. Se entiende como una secuencia de experimentos de Bernoulli independientes y también puede plantearse como una secuencia de r casos independientes de la distribución Geométrica que se explicará más adelante.

Para obtener la función de probabilidad y la función de distribución de la distribución Binomial Negativa se aplica la siguiente fórmula:

$$f_x(x) = \binom{x+r-1}{r-1} p^r (1-p)^x; \text{ siendo } x = 0, 1, 2, 3 \dots \quad \text{[Fórmula 7]}$$

$$F_x(x) = \sum_{i=0}^x \binom{i+r-1}{i} p^r (1-p)^i I_{(0,1,2,\dots)}(i) \quad \text{[Fórmula 8]}$$

Los estadísticos de la distribución se calculan de la manera que sigue a continuación:

$$E(X) = \frac{r(1-p)}{p}$$

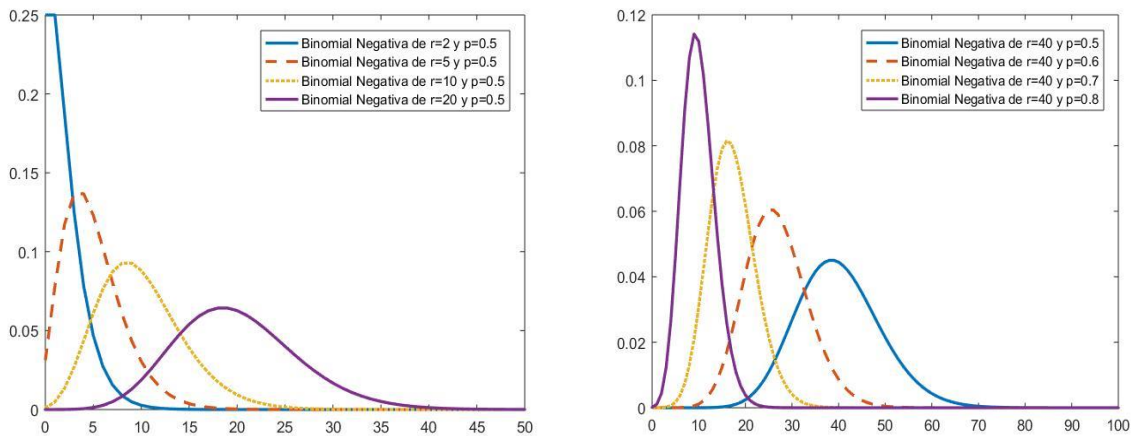
[Fórmula 9]

$$Var(X) = \frac{r(1-p)}{p^2}$$

[Fórmula 10]

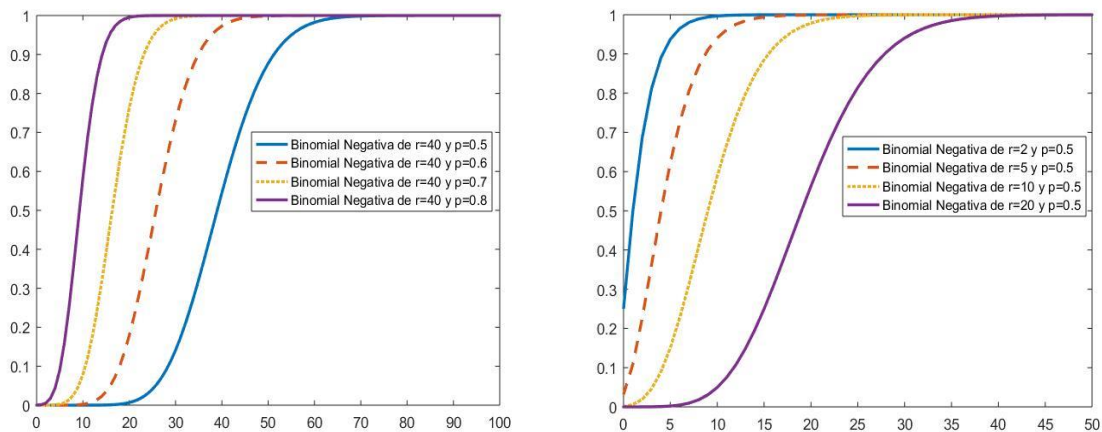
A continuación, se muestran algunos ejemplos gráficos del comportamiento de la distribución Binomial Negativa alternando los valores de los parámetros r y p .

Gráfico 3. Función de probabilidad de la distribución Binomial Negativa (r,p).



Fuente: Elaboración propia.

Gráfico 4. Función de distribución de la distribución Binomial Negativa (r,p).



Fuente: Elaboración propia.

Distribución Geométrica

La distribución Geométrica es una distribución discreta uniparamétrica caracterizada por el parámetro p para su modelización. Esta distribución es utilizada en procesos en los que el interés se centra en calcular la probabilidad de que un éxito se produzca por única vez en el último ensayo del experimento realizado.

El parámetro p representa la probabilidad de alcanzar el evento de éxito en cada fallo, por lo que su valor oscilará entre 0 y 1. Es un caso particular de la distribución Binomial Negativa estudiada anteriormente cuando r toma el valor 1, es decir, siendo este el supuesto en el que sólo interesa alcanzar el evento de éxito una vez.

Para obtener la función de probabilidad y la función de distribución de la distribución Geométrica, se recurre a la siguiente fórmula:

$$f_x(x) = p(1 - p)^x; \text{ siendo } x = 0, 1, 2, 3 \dots \quad [\text{Fórmula 11}]$$

$$F_x(x) = 1 - (1 - p)^{x+1}; \text{ siendo } x = 0, 1, 2, 3 \dots \quad [\text{Fórmula 12}]$$

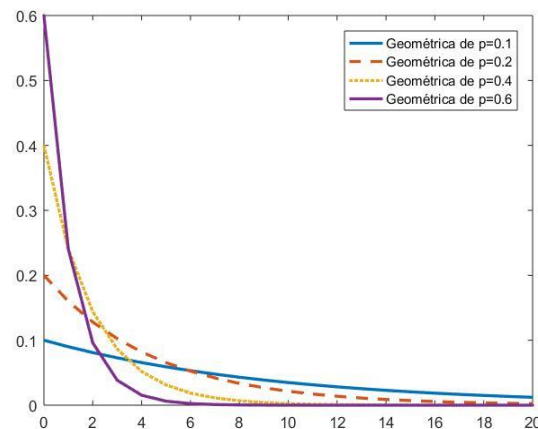
Los estadísticos de la distribución se calculan de la manera que sigue a continuación:

$$E(X) = \frac{(1 - p)}{p} \quad [\text{Fórmula 13}]$$

$$Var(X) = \frac{(1 - p)}{p^2} \quad [\text{Fórmula 14}]$$

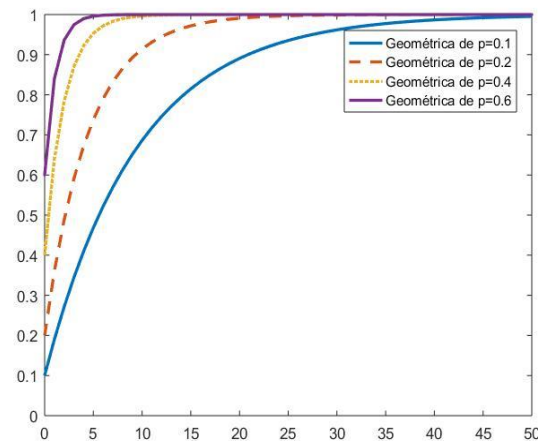
A continuación, se muestran algunos ejemplos gráficos del comportamiento de la distribución Geométrica alternando los valores del parámetro p .

Gráfico 5. Función de probabilidad de la distribución Geométrica (p).



Fuente: Elaboración propia.

Gráfico 6. Función de distribución de la distribución Geométrica (p).



Fuente: Elaboración propia.

Distribución Poisson

La distribución de Poisson es una distribución uniparamétrica, siendo λ el parámetro que define esta distribución de probabilidad discreta. Es una distribución de especial interés en el ámbito actuarial, ya que se aplica en casos de modelización cuyo objetivo es determinar el número de eventos acontecidos en un intervalo acotado de tiempo o de espacio.

El parámetro de la distribución de Poisson, λ , tiene como restricción ser un número no negativo. Este parámetro tiene la singularidad de representar tanto media como varianza de la distribución.

Para obtener la función de probabilidad y la función de distribución de la distribución Poisson se aplica la siguiente fórmula:

$$f_x(x) = \frac{\lambda^x e^{-\lambda}}{x!} \quad \text{[Fórmula 15]}$$

$$F_x(x) = e^{-\lambda} \sum_{i=0}^x \frac{\lambda^i}{i!} \quad \text{[Fórmula 16]}$$

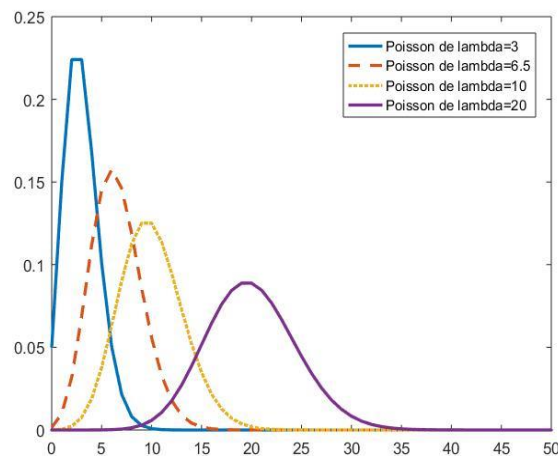
Los estadísticos de la distribución de Poisson cumplen la siguiente característica particular:

$$E(X) = \lambda \quad \text{[Fórmula 17]}$$

$$Var(X) = \lambda \quad \text{[Fórmula 18]}$$

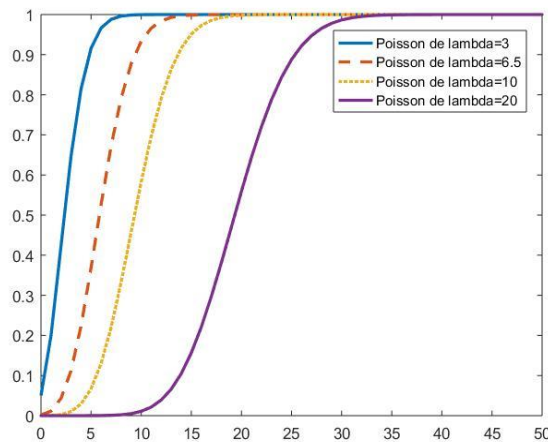
A continuación, se muestran algunos ejemplos gráficos del comportamiento de la distribución de Poisson alternando los valores del parámetro λ .

Gráfico 7. Función de probabilidad de la distribución Poisson (λ).



Fuente: Elaboración propia.

Gráfico 8. Función de distribución de la distribución Poisson (λ).



Fuente: Elaboración propia.

3.1.2. Distribuciones discretas de la clase (a, b, 1)

Existe otra familia de distribuciones similares a las anteriormente expuestas y que se reconocen con el nombre de distribuciones de la clase (a, b, 1). Al igual que las anteriores, cumplen la siguiente fórmula recursiva:

$$f(x) = \left(a + \frac{b}{x}\right) f(x - 1); \text{ siendo } x = 2, 3, 4, 5 \dots$$

[Fórmula 19]

La diferencia de la fórmula aplicada para esta clase y la anterior es el punto donde empieza. Para la clase (a, b, 0) bastaba con calcular el valor de la función en 0 y utilizar la fórmula desde 1, mientras que, en este segundo caso, la fórmula empieza a aplicarse desde 2, por lo que se necesitaría el valor de la función en 1.

Mediante la utilización de las distribuciones de la clase (a,b,1) se consiguen nuevas distribuciones con una probabilidad en el punto 0 específica, es decir, puede determinarse la probabilidad que se tomará en el punto 0 de la distribución. Como consecuencia, se modifican las sucesivas probabilidades de la distribución, pero se mantiene la forma y características de la distribución de probabilidad que ha sido utilizada como base (Binomial, Binomial Negativa, Geométrica o Poisson).



Una vez determinado el valor con el que se quiere modificar la probabilidad en el punto o de la distribución, basta con multiplicar la siguiente constante a los valores de las probabilidades siguientes a o :

$$c = \frac{1 - f_x^M(o)}{1 - f_x(o)}$$

[Fórmula 20]

Donde:

$f_x^M(o)$ es la probabilidad asignada que tomará la distribución en o .

$f_x(o)$ es la probabilidad en o de la distribución original.

El motivo de incluir esta familia de distribuciones de la clase $(a,b,1)$ es disponer de una gama de distribuciones discretas que puedan ajustarse mejor a las frecuencias de siniestros de cada grupo homogéneo. En consecuencia, dotará el presente estudio de mayor flexibilidad y permitirá realizar ajustes más exactos al comportamiento presentado.

Para el caso de las distribuciones modificadas en o (aquellas distribuciones seguidas de la terminación “ZM” en el resultado del aplicativo), se fijará un valor al inicio de la ejecución del aplicativo que será la probabilidad que las distribuciones evaluadas tomarán en dicho punto. Para las distribuciones truncadas en el punto o (aquellas distribuciones seguidas de la terminación “ZT”), el aplicativo realiza el cálculo directamente para evaluar el posible ajuste de la distribución truncada a los datos de frecuencias correspondientes al clúster analizado.

Se resume en la *Tabla 2* los valores que tomarían los parámetros “ a ” y “ b ” para las distribuciones de la clase $(a, b, 1)$.



Tabla 2. Resumen de los parámetros “a” y “b” y cálculo de $f_x(0)$ de las distribuciones discretas de la clase $(a,b,1)$.

Distribuciones	Parámetro “a”	Parámetro “b”	$f_x(0)$
Binomial ZM Binomial ZT	$\frac{-p}{(1-p)}$	$\frac{(n+1)p}{(1-p)}$	Valor fijado 0
Binomial Negativa ZM Binomial Negativa ZT	$\frac{\beta}{1+\beta}$	$(r-1)\frac{\beta}{1+\beta}$	Valor fijado 0
Geométrica ZM Geométrica ZT	$\frac{\beta}{1+\beta}$	0	Valor fijado 0
Poisson ZM Poisson ZT	0	λ	Valor fijado 0

Fuente: Elaboración propia a partir de la información obtenida del libro “Loss Models: From data to decisions” (2012).

3.2. Estudio de severidades de siniestros

La severidad refleja el coste de los siniestros que el tomador del seguro presenta y que la compañía debe pagar al asegurado para satisfacer la pérdida sufrida como consecuencia del acaecimiento del evento asegurado. Esta variable, a diferencia con la anterior, presenta características de una variable aleatoria continua, por lo que se ha seleccionado un abanico de distribuciones para ajustar los datos de severidad presentados en cada clúster: Gamma, Exponencial, Pareto Generalizada, Inversa Gaussiana, Logistic, LogLogistic, LogNormal, Normal y Weibull.



Distribución Gamma

La distribución Gamma es una distribución continua de dos parámetros, α y β . Entre sus aplicaciones en el ramo No Vida, puede destacarse su utilización para la modelización del tiempo que transcurre hasta que sucede un determinado número de veces un evento.

Los parámetros α y β de la distribución Gamma tomarán como valores números superiores a 0. Es un caso de especial interés para la modelización actuarial, puesto que bajo distintas condiciones en sus parámetros da lugar a nuevas distribuciones de probabilidad, como es el caso de la distribución Exponencial que se estudiará más adelante.

La expresión de la función de densidad y la función de distribución de la distribución Gamma es la siguiente:

$$f_x(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} e^{-\frac{x}{\beta}} x^{\alpha-1};$$

siendo $\Gamma(x)$ la función Gamma de Euler

[Fórmula 21]

$$F_x(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^x t^{\alpha-1} e^{-\frac{t}{\beta}} dt;$$

siendo $\Gamma(x)$ la función Gamma de Euler

[Fórmula 22]

Los estadísticos de esta distribución continua Gamma son los siguientes:

$$E(X) = \alpha\beta$$

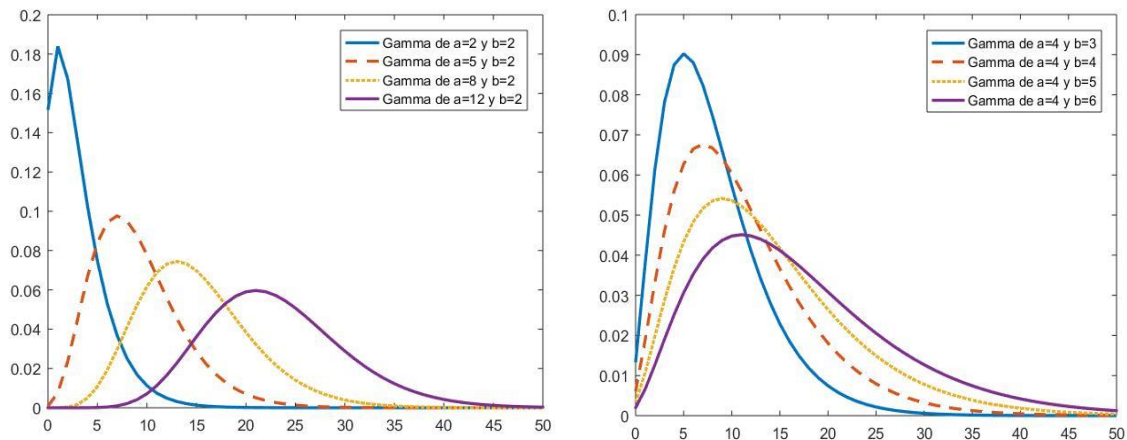
[Fórmula 23]

$$Var(X) = \alpha\beta^2$$

[Fórmula 24]

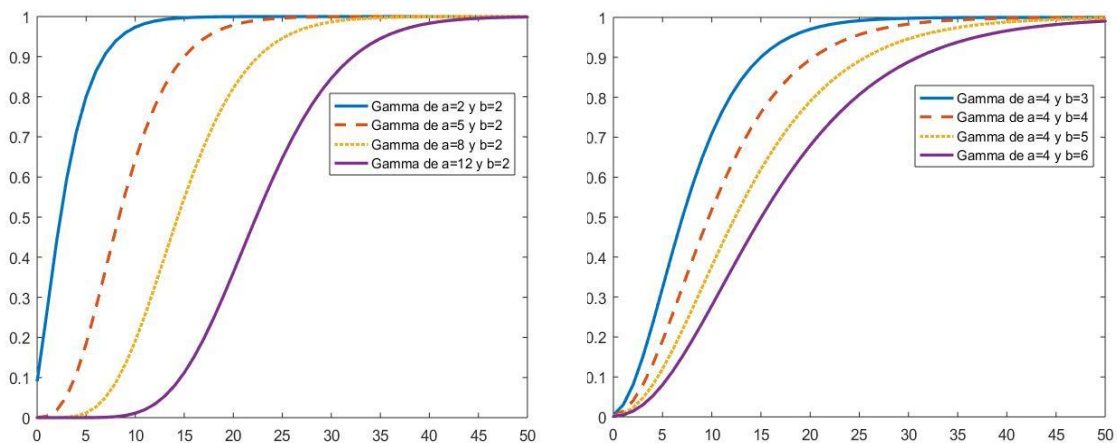
A continuación, se muestran algunos ejemplos gráficos del comportamiento de la distribución Gamma alternando los valores de los parámetros α y β .

Gráfico 9. Función de probabilidad de la distribución Gamma (α, β).



Fuente: Elaboración propia.

Gráfico 10. Función de distribución de la distribución Gamma (α, β).



Fuente: Elaboración propia.

Distribución Exponencial

La distribución Exponencial es un caso particular de la distribución Gamma explicada anteriormente. Está caracterizada por los mismos parámetros α y β , pero tomando α el valor 1. Es la versión continua de la distribución Geométrica y suele aplicarse a estudios donde interesa conocer cuánto tiempo transcurre hasta el acaecimiento de un evento, igual que la Gamma.

Al tratarse de un caso particular de la distribución Gamma, el parámetro β de la distribución Exponencial tomará también valores números superiores a 0.

La expresión de la función de densidad y la función de distribución de la distribución Exponencial es la siguiente:

$$f_x(x) = \frac{1}{\beta} e^{-\frac{x}{\beta}}$$

[Fórmula 25]

$$F_x(x) = \frac{1}{\beta} \int_0^x e^{-\frac{t}{\beta}} dt$$

[Fórmula 26]

Los estadísticos de esta distribución continua Exponencial son los siguientes:

$$E(X) = \beta$$

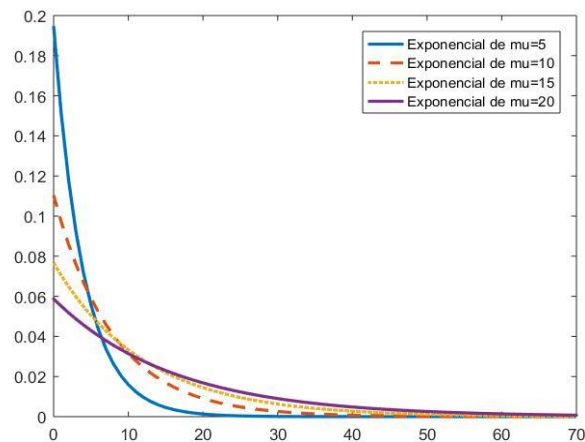
[Fórmula 27]

$$Var(X) = \beta^2$$

[Fórmula 28]

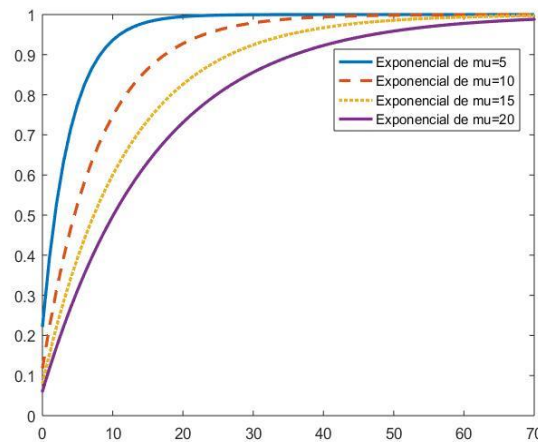
A continuación, se muestran algunos ejemplos gráficos del comportamiento de la distribución Exponencial alternando los valores de los parámetros α y β .

Gráfico 11. Función de probabilidad de la distribución Exponencial (μ).



Fuente: Elaboración propia.

Gráfico 12. Función de distribución de la distribución Exponencial (μ).



Fuente: Elaboración propia.

Distribución Pareto Generalizada

La distribución Pareto Generalizada es una distribución continua de tres parámetros: el parámetro de forma k , definido como un escalar; el parámetro de escala σ , especificado como un escalar no negativo, y el parámetro de umbral θ , definido como un escalar. Es una distribución que suele asociarse con la modelización de siniestros catastróficos debido a que cuenta como característica principal el tener una cola pesada, lo que implica acumular una masa de probabilidad considerable en la cola de la distribución.

La expresión de la función de densidad y la función de distribución de la distribución Pareto Generalizada es la siguiente, según la casuística:

Para el caso de que el parámetro θ de la distribución sea inferior a x , k sea mayor que 0, o $\theta < x < \theta + \frac{\sigma}{k}$, siendo k inferior a 0.

$$f_x(x) = \left(\frac{1}{\sigma}\right) \left(1 + k \frac{(x - \theta)}{\sigma}\right)^{-1 - \frac{1}{k}}$$

[Fórmula 29]

$$F_x(x) = 1 - \left(1 + k \frac{(x - \theta)}{\sigma}\right)^{-\frac{1}{k}}$$

[Fórmula 30]

En el caso en el que el parámetro de forma k tome el valor 0, la función de densidad será la siguiente:

$$f_x(x) = \left(\frac{1}{\sigma}\right) e^{-\frac{(x-\theta)}{\sigma}}$$

[Fórmula 31]

$$F_x(x) = 1 - e^{-\frac{(x-\theta)}{\sigma}}$$

[Fórmula 32]

Los estadísticos de esta distribución continua Pareto Generalizada son los siguientes:

$$E(X) = \theta + \frac{\sigma}{1-k}; \text{ siendo } k < 1$$

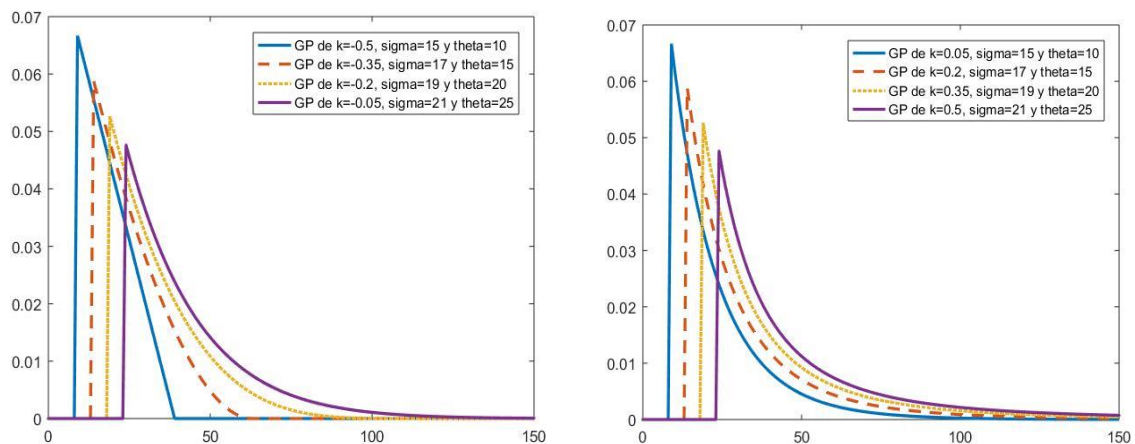
[Fórmula 33]

$$Var(X) = \frac{\sigma^2}{(1-2k)(1-k)^2}; \text{ siendo } k > 0,5$$

[Fórmula 34]

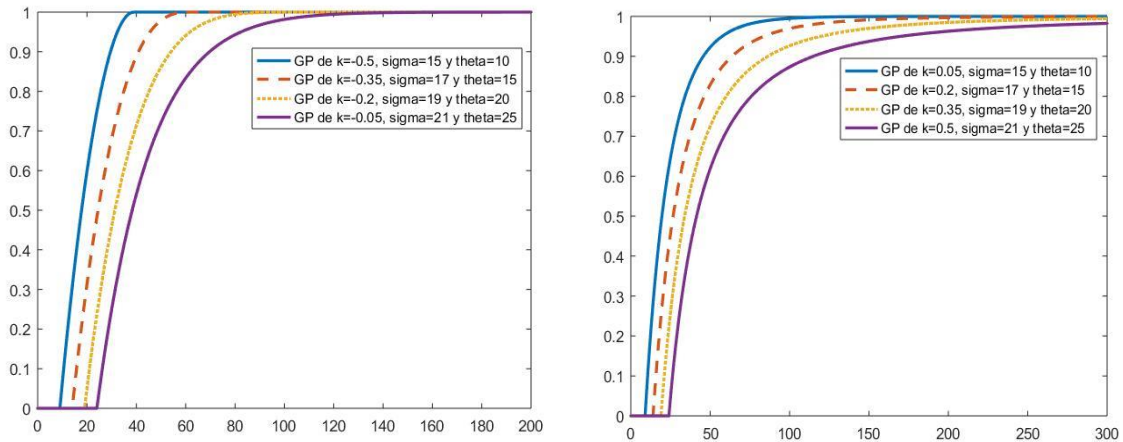
A continuación, se muestran algunos ejemplos gráficos del comportamiento de la distribución Pareto Generalizada alternando los valores de los parámetros k , σ y θ .

Gráfico 13. Función de probabilidad de la distribución Pareto Generalizada (k , σ , θ).



Fuente: Elaboración propia.

Gráfico 14. Función de distribución de la distribución Pareto Generalizada (k, σ, θ).



Fuente: Elaboración propia.

Distribución Inversa Gaussiana

Esta distribución de probabilidad continua, también conocida como la distribución de Wald, está definida por los parámetros μ y λ . Es una distribución adecuada cuando el conjunto de datos a modelizar presenta una asimetría positiva. Además, esta distribución posee particularidades similares a la distribución Normal, que se estudiará a continuación. Esta característica hace de esta distribución un caso interesante para la modelización de datos tanto en el mundo actuarial como en el mundo de las finanzas para el estudio de la evolución del precio de activos financieros.

La distribución continua Inversa Gaussiana está definida por los parámetros μ y λ , que hacen referencia a la escala y la forma de esta distribución de probabilidad. Ambos parámetros son definidos como escalares no negativos.

La expresión de la función de densidad y la función de distribución de la distribución Inversa Gaussiana es la siguiente:

$$f_x(x) = \sqrt{\frac{\lambda}{2\pi x^3}} e^{-\frac{\lambda(x-\mu)^2}{2\mu^2 x}}$$

[Fórmula 35]

$$F_x(x) = \Phi\left(\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu} - 1\right)\right) + e^{\left(\frac{2\lambda}{\mu}\right)}\Phi\left(-\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu} + 1\right)\right);$$

siendo Φ la función de distribución Normal

[Fórmula 36]

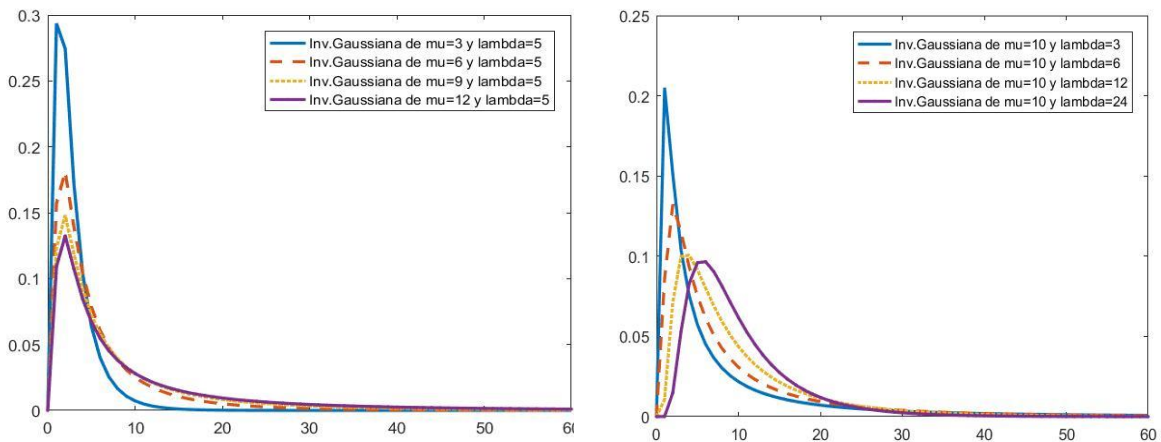
Los estadísticos para la distribución continua Inversa Gaussiana son los siguientes:

$$E(X) = \mu \tag{Fórmula 37}$$

$$Var(X) = \frac{\mu^3}{\lambda} \tag{Fórmula 38}$$

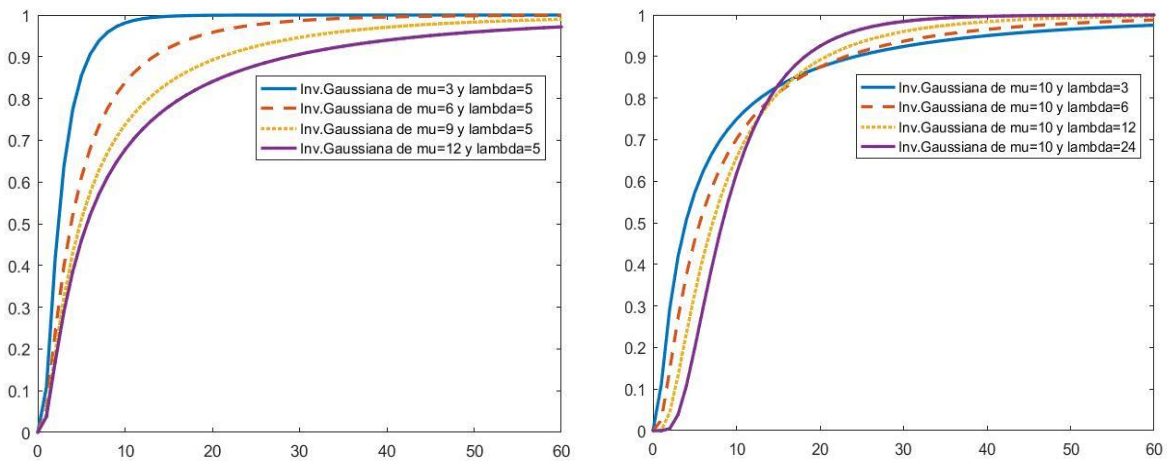
A continuación, se muestran algunos ejemplos gráficos del comportamiento de la distribución biparamétrica Inversa Gaussiana alternando los valores de los parámetros μ y λ .

Gráfico 15. Función de probabilidad de la distribución Inversa Gaussiana (μ, λ).



Fuente: Elaboración propia.

Gráfico 16. Función de distribución de la distribución Inversa Gaussiana (μ, λ).



Fuente: Elaboración propia.

Distribución Logistic

La distribución Logistic es una distribución biparamétrica definida por μ y σ . Es un modelo que aparece con frecuencia en estudios de procesos de crecimiento y difusión respecto de un estado inicial (propagación de enfermedades, venta de productos nuevos, mortalidad...). Como se puede apreciar en el *Gráfico 17*, presenta un comportamiento similar al de la distribución Normal, pero tiene como diferencia que las colas de la distribución son más pesadas.

Los parámetros que modelizan el comportamiento de esta distribución son μ , la media, cuyo soporte se extiende entre $-\infty < x < \infty$, y σ , que es el parámetro de escala de la distribución y que puede tomar como valor cualquier número igual o superior a 0.

La expresión de la función de densidad y la función de distribución de la distribución Logistic es la siguiente:

$$f_x(x) = \frac{e\left(-\frac{x-\mu}{\sigma}\right)}{\sigma \left(1 + e\left(-\frac{x-\mu}{\sigma}\right)\right)^2} \quad \text{[Fórmula 39]}$$

$$F_x(x) = \frac{1}{1 + e^{-\frac{x-\mu}{\sigma}}} \quad \text{[Fórmula 40]}$$

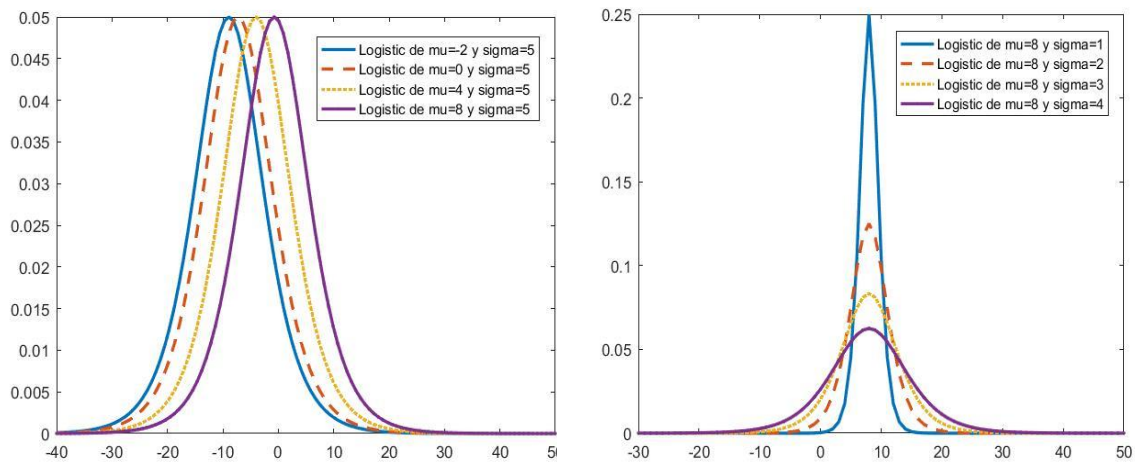
Los estadísticos para la distribución continua Logistic son los siguientes:

$$E(X) = \mu \quad \text{[Fórmula 41]}$$

$$Var(X) = \frac{\pi^2}{3} \sigma^2 \quad \text{[Fórmula 42]}$$

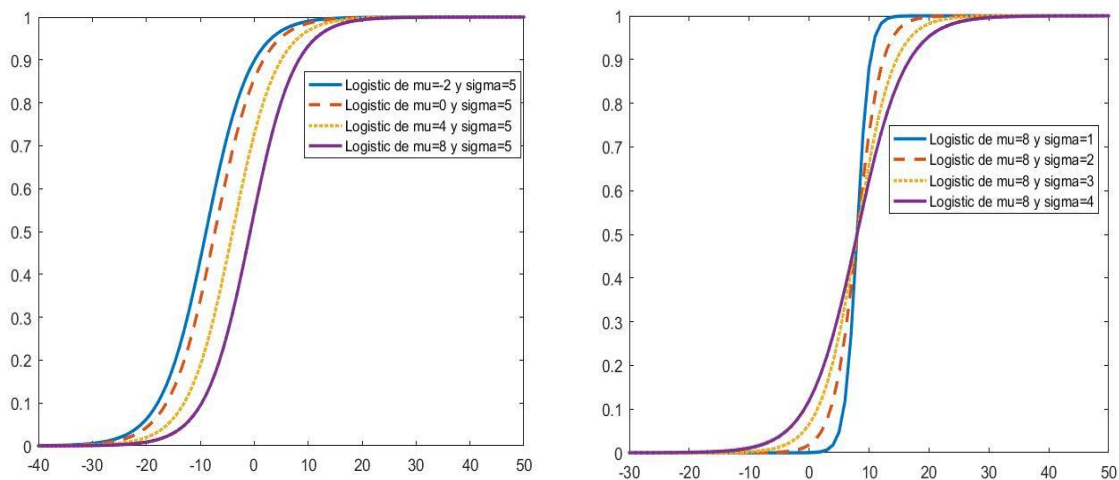
A continuación, se muestran algunos ejemplos gráficos del comportamiento de la distribución Logistic alternando los valores de los parámetros μ y σ .

Gráfico 17. Función de probabilidad de la distribución Logistic (μ, σ).



Fuente: Elaboración propia.

Gráfico 18. Función de distribución de la distribución Logistic (μ, σ).



Fuente: Elaboración propia.

Distribución LogLogistic

La distribución LogLogistic se caracteriza también por los parámetros μ y σ . Sus aplicaciones prácticas suelen estar enfocadas al análisis de la vida útil de determinados objetos o mecanismos. En el mundo de investigaciones económicas suele denominarse como distribución de Fisk. La principal característica de esta distribución, tal como puede apreciarse en el *Gráfico 20*, es la cola, que para esta distribución resulta ser muy pesada, es decir, acumula mucha probabilidad.

Los parámetros μ y σ , que al igual que en el caso anterior representan la media y la escala. Al tratarse de una distribución logarítmica no tomará valores iguales o inferiores a 0.

La expresión de la función de densidad y la función de distribución de la distribución LogLogistic es la siguiente:

$$f_x(x) = \frac{\sigma\mu(\sigma x)^{\mu-1}}{(1 + (\sigma x)^\mu)^2} \quad \text{[Fórmula 43]}$$

$$F_x(x) = \frac{(\sigma x)^\mu}{1 + (\sigma x)^\mu} \quad \text{[Fórmula 44]}$$

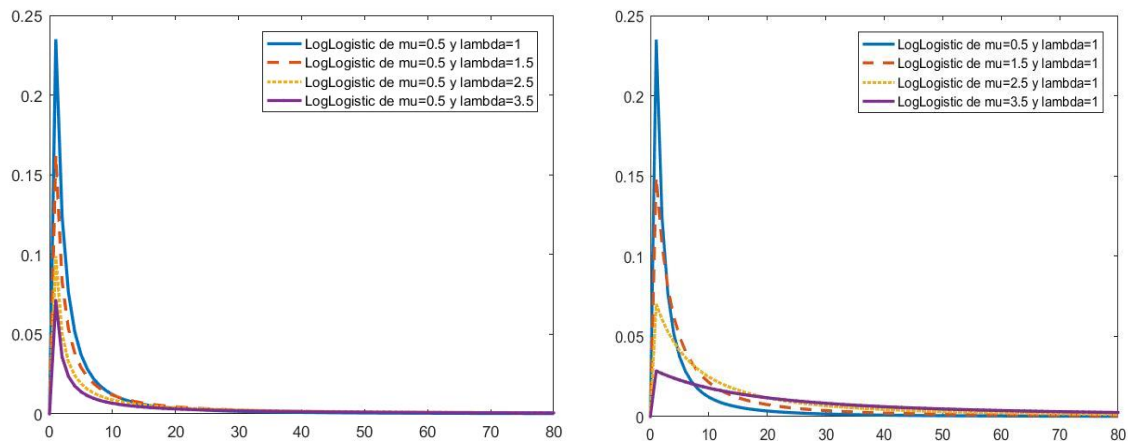
Los estadísticos para la distribución continua LogLogistic son:

$$E(X) = \frac{\pi}{\mu\sigma \left(\text{sen} \left(\frac{\pi}{\mu} \right) \right)} \quad \text{[Fórmula 45]}$$

$$\text{Var}(X) = \frac{\pi \left(2\mu \left(1 - \cos \left(\frac{\pi}{\sigma} \right)^2 \right) + \pi \text{sen} \left(\frac{\pi(\mu + 2)}{\mu} \right) \right)}{\left(\text{sen} \left(\frac{\pi(\mu + 2)}{\mu} \right) \right) \left(\left(\cos \left(\frac{\pi}{\mu} \right) \right)^2 - 1 \right) (\mu\sigma)^2} \quad \text{[Fórmula 46]}$$

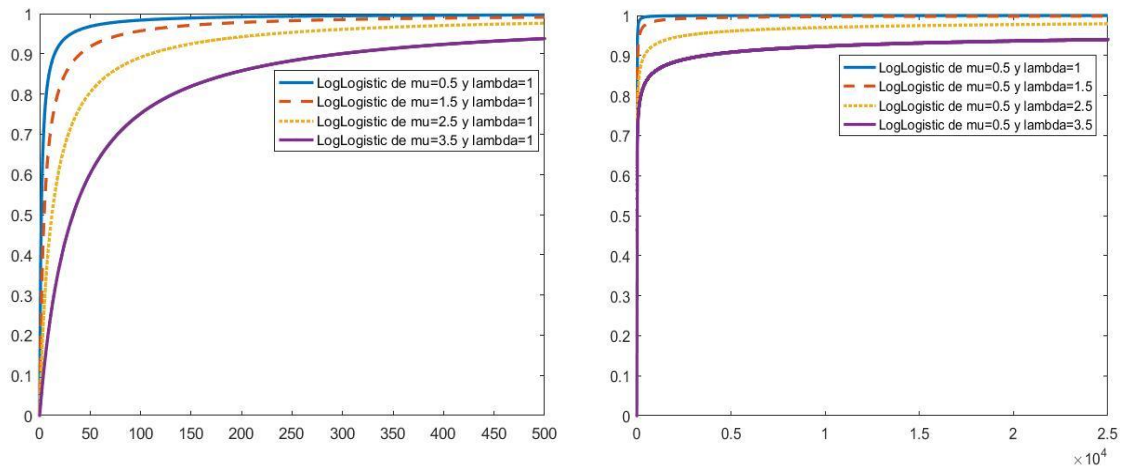
A continuación, se muestran algunos ejemplos gráficos del comportamiento de la distribución LogLogistic alternando los valores de los parámetros μ y σ .

Gráfico 19. Función de probabilidad de la distribución LogLogistic (μ, σ).



Fuente: Elaboración propia.

Gráfico 20. Función de distribución de la distribución LogLogistic (μ, σ)



Fuente: Elaboración propia.

Distribución LogNormal

La distribución LogNormal es una distribución biparamétrica que deriva de la Normal y que se caracteriza por los parámetros μ y σ . Esta distribución surge cuando el logaritmo de una variable aleatoria que toma valores positivos se distribuye según una Normal. Es frecuente aplicar esta distribución para modelizar la vida útil de diversos instrumentos e instalaciones.

En cuanto a los parámetros que definen esta distribución de probabilidad, el parámetro μ representa la media y el parámetro σ la desviación típica. Tal y como puede observarse en el *Gráfico 22*, manteniendo una media constante e incrementando la desviación, la cola de la distribución tenderá a ser más pesada.

La expresión de la función de densidad y la función de distribución de la distribución LogNormal es la siguiente:

$$f_x(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}$$

[Fórmula 47]

$$F_x(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_0^x \frac{e^{-\frac{(\ln t - \mu)^2}{2\sigma^2}}}{t} dt$$

[Fórmula 48]

Los estadísticos para la distribución continua LogNormal son los siguientes:

$$E(X) = e^{\left(\mu + \frac{\sigma^2}{2}\right)}$$

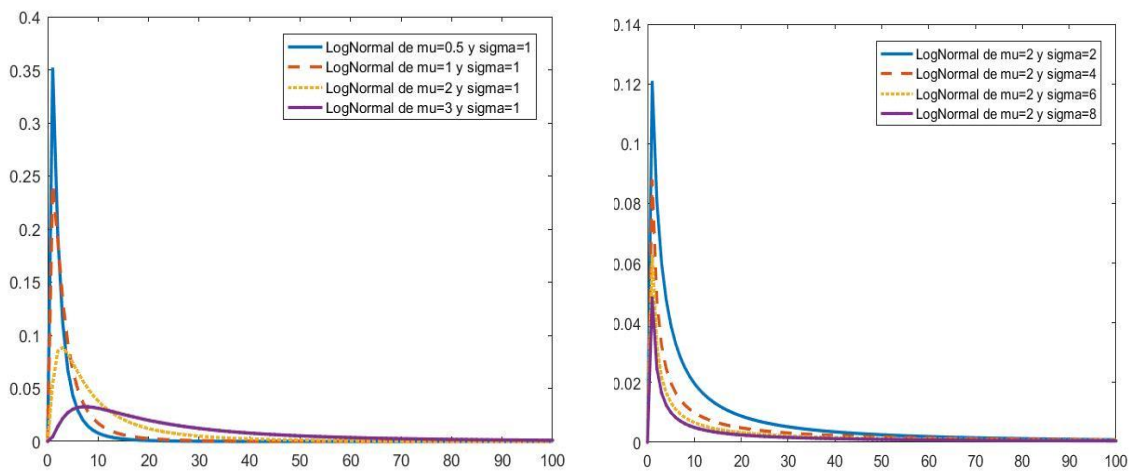
[Fórmula 49]

$$Var(X) = (e^{\sigma^2} - 1)e^{2\mu + \sigma^2}$$

[Fórmula 50]

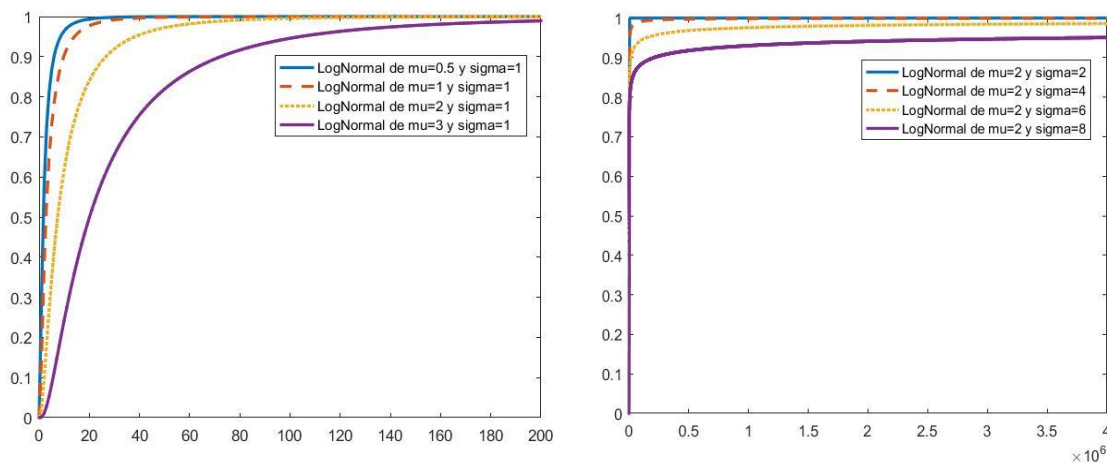
A continuación, se muestran algunos ejemplos gráficos del comportamiento de la distribución LogNormal alternando los valores de los parámetros μ y σ .

Gráfico 21. Función de probabilidad de la distribución LogNormal (μ, σ).



Fuente: Elaboración propia.

Gráfico 22. Función de distribución de la distribución LogNormal (μ, σ).



Fuente: Elaboración propia.

Distribución Normal

La distribución Normal es la distribución de probabilidad más conocida y una de las más importantes. Su principal característica es la simetría y que las medidas de posición central media, moda y mediana coinciden. Además, cualquier transformación lineal de una distribución Normal tiene como resultado otra distribución Normal.

En cuanto a sus parámetros, el parámetro μ representa la media de la distribución y el parámetro σ la desviación típica, al igual que en el caso de la LogNormal.

La expresión de la función de densidad y la función de distribución de la distribución Normal es la siguiente:

$$f_x(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

[Fórmula 51]

$$F_x(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$$

[Fórmula 52]

Los estadísticos para la distribución continua Normal son los siguientes:

$$E(X) = \mu$$

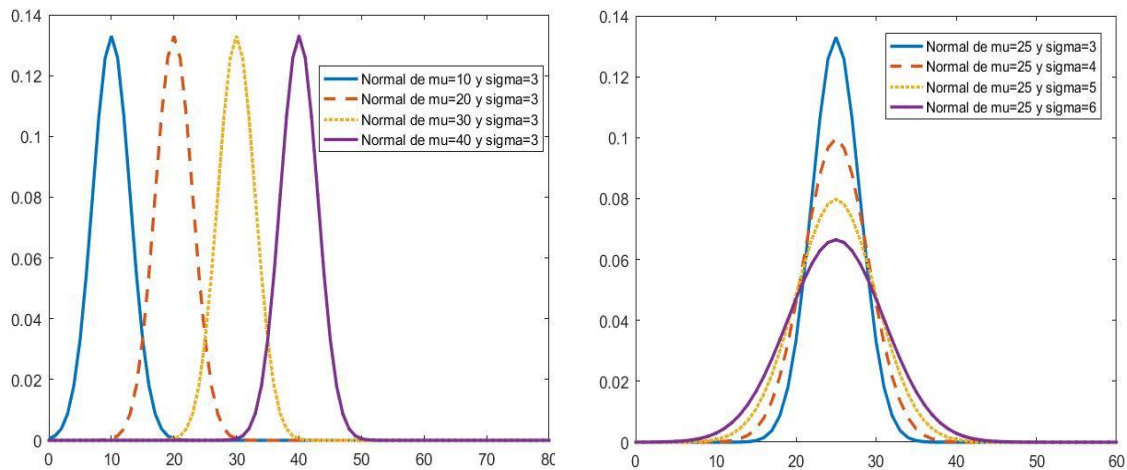
[Fórmula 53]

$$Var(X) = \sigma^2$$

[Fórmula 54]

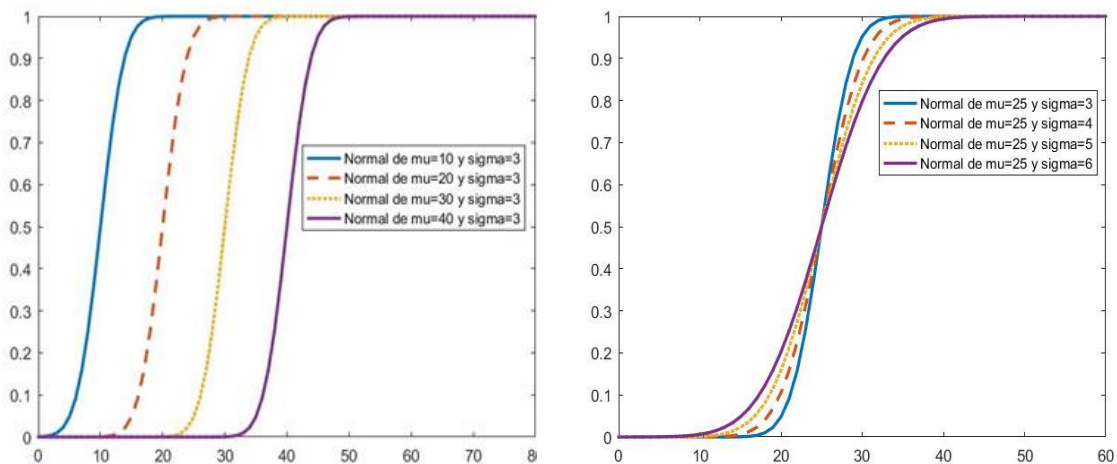
A continuación, se muestran algunos ejemplos gráficos del comportamiento de la distribución Normal alternando los valores de los parámetros μ y σ .

Gráfico 23. Función de probabilidad de la distribución Normal (μ, σ).



Fuente: Elaboración propia.

Gráfico 24. Función de distribución de la distribución Normal (μ, σ).



Fuente: Elaboración propia.

Distribución Weibull

La distribución continua Weibull es una está definida por los parámetros A y B y que, al igual que alguna de las distribuciones estudiadas anteriormente, suele emplearse para analizar el tiempo de vida útil de equipos o la resistencia de materiales, entre otras aplicaciones.

En cuanto a sus parámetros, A es un parámetro de escala y B es el parámetro de forma, tomando ambos parámetros valores superiores a 0. En el caso de que el parámetro B sea igual a 1, esta distribución Weibull será una distribución Exponencial.

La expresión de la función de densidad y la función de distribución de la distribución Weibull es la siguiente:

$$f_x(x) = \frac{B}{A} \left(\frac{x}{A}\right)^{B-1} e^{-\left(\frac{x}{A}\right)^B}$$

[Fórmula 55]

$$F_x(x) = \int_0^x BA^{-B} t^{B-1} e^{-\left(\frac{t}{A}\right)^B} dt$$

[Fórmula 56]

Los estadísticos para la distribución continua Weibull son los siguientes:

$$E(X) = A\Gamma(1 + B^{-1});$$

siendo $\Gamma(x)$ la función Gamma de Euler

[Fórmula 57]

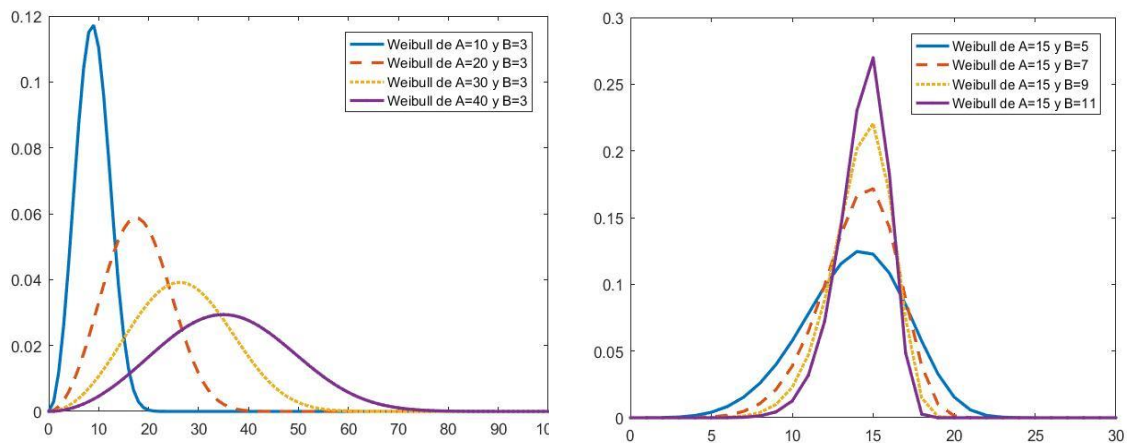
$$Var(X) = A^2[\Gamma(1 + 2B^{-1}) - \Gamma(1 + B^{-1})^2];$$

siendo $\Gamma(x)$ la función Gamma de Euler

[Fórmula 58]

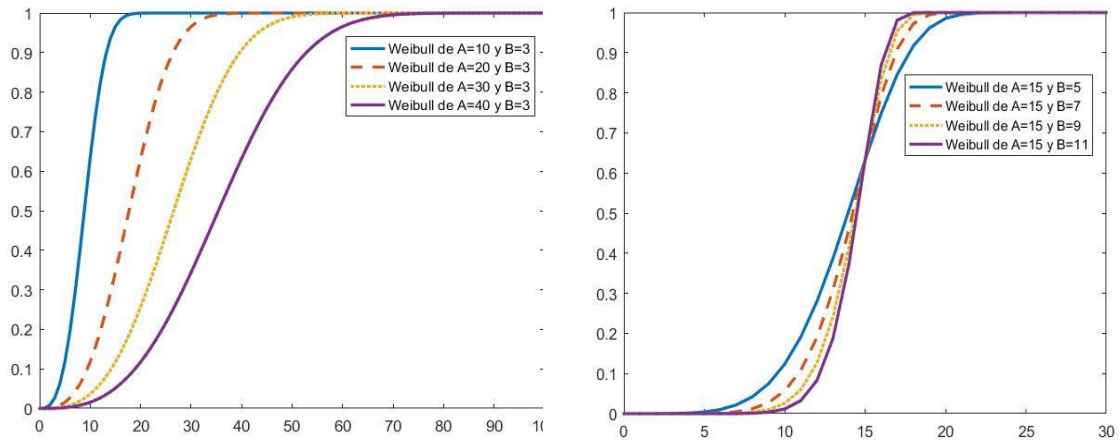
A continuación, se muestran algunos ejemplos gráficos del comportamiento de la distribución Weibull alternando los valores de los parámetros A y B .

Gráfico 25. Función de probabilidad de la distribución Weibull (A, B).



Fuente: Elaboración propia.

Gráfico 26. Función de distribución de la distribución Weibull (A, B).



Fuente: Elaboración propia.



4. Elección de la distribución de probabilidad de frecuencias y severidades

La selección de la distribución de probabilidad que mejor se ajuste a las frecuencias y severidades de siniestros presentados por los clústeres es un proceso crucial. De ella dependerá lo adecuada que resulte la construcción de la distribución de pérdidas agregadas por Panjer y el proceso de tarificación posterior, objetivo principal del presente TFM.

La importancia reside en que, como se ha explicado en el epígrafe anterior, las distribuciones tienen comportamientos distintos que afectan directamente a los estadísticos media y varianza y, por extensión, a la medición de los riesgos. Esto supone que seleccionar un modelo adecuado reflejará de manera certera el comportamiento de los asegurados de cada clúster e influirá directamente en la tarifa a cobrar por el seguro.

A continuación, se explica el proceso de cómo se determinará la distribución de probabilidad que mejor representa las variables frecuencias y severidad. En primer lugar, mediante el Método de Máxima Verosimilitud se estiman los parámetros que puede seguir una distribución de probabilidad en base a las observaciones de frecuencias y severidades. Seguidamente, se efectúa el Test Chi-Cuadrado para comprobar si efectivamente la distribución a la cual se han estimado los parámetros por Máxima Verosimilitud puede ajustarse a los datos observados de frecuencias y severidades. Finalmente, se comprueba el p-Valor para evaluar con qué seguridad se puede aceptar la hipótesis arrojada por el Test Chi-Cuadrado.

4.1. Estimación de parámetros de la distribución por Máxima Verosimilitud

El Método de Máxima Verosimilitud es una forma de calcular cómo de creíble es el estimador de un parámetro en base a la información que contiene la muestra concreta, asumiendo que tiende a suceder lo más probable. En esta línea, siguiendo la distribución de probabilidad poblacional se construye la función de verosimilitud para determinar cuál es el valor que tomará el parámetro el cual trata de estimarse.

Existen otras formas de estimar parámetros, como el Método de los Momentos. Este método es más sencillo, puesto que se basa en establecer la igualdad entre los momentos muestrales respecto del origen y los poblacionales. Pero en esa sencillez reside su principal desventaja, dado que para definir un estimador sólo se centra en la información facilitada por los momentos muestrales y poblacionales en lugar de analizar toda la información de la muestra y poblacional.



Frente al Método de los Momentos, el Método de Máxima Verosimilitud mejora los resultados alcanzados debido a que presenta las siguientes propiedades:

1. El Estimador de Máxima Verosimilitud está dentro del espacio muestral.
2. El Estimador de Máxima Verosimilitud se distribuye asintóticamente según la Normal.
3. El Estimador de Máxima Verosimilitud es asintóticamente insesgado.
4. El Estimador de Máxima Verosimilitud es consistente, es decir, cuando la muestra es grande, el valor del estimador tiende al valor del parámetro.
5. El Estimador de Máxima Verosimilitud es asintóticamente eficiente (tiende a mínima varianza).

4.1.1. Método de Máxima Verosimilitud para casos discretos

Para el caso discreto, la función de verosimilitud se obtiene como probabilidad de la muestra como función del parámetro, es decir, siendo la variable de la función el parámetro. De este modo, teniendo una muestra de la variable x y la función de probabilidad poblacional $P(\xi=x, \theta)$ se consigue una función de verosimilitud definida como proporcional a la probabilidad de aparición de esa muestra de la variable condicionada al parámetro.

Matemáticamente, se representa la función de verosimilitud “ L ” de la siguiente manera:

$$L(x_1, x_2, x_3, \dots, x_n, \theta) = \prod_{i=1}^n P(\xi = x_i, \theta)$$

[Fórmula 59]

Una vez obtenida la función de verosimilitud, el Método consiste en obtener el valor del parámetro que la maximiza. En este sentido, a fin de simplificar los cálculos, se aplica el logaritmo neperiano de la función de verosimilitud, lo que se denomina función soporte. La expresión matemática de esta idea es:

$$\ln L(\hat{\theta}; X) = \underbrace{\max}_{\theta \in \Theta} \ln L(\theta; X)$$

[Fórmula 60]

Donde:

$\hat{\theta}$ es el estimador del parámetro obtenido por máxima verosimilitud.

4.1.2. Método de Máxima Verosimilitud para casos continuos

Para las continuas, el cálculo de la función de verosimilitud se obtiene de manera similar al caso discreto, pero con la densidad de la muestra asumida como función del parámetro. De este modo, la función de densidad poblacional $f_{\xi}(x, \theta)$ se obtiene igual que en el caso discreto: una función de verosimilitud definida como proporcional a la probabilidad de aparición de esa muestra de la variable condicionada al parámetro.

Matemáticamente, se representa la función de verosimilitud “L” de la siguiente manera:

$$L(x_1, x_2, x_3, \dots, x_n, \theta) = \prod_{i=1}^n f_{\xi}(x_i, \theta)$$

[Fórmula 61]

La expresión matemática de la función de Máxima Verosimilitud para el caso continuo coincide con la *Fórmula 60*.

En conclusión, y a modo de resumen, el Método de Máxima Verosimilitud consiste en obtener una muestra del experimento y fijar una función de probabilidad conocida. A continuación, se construye la función de verosimilitud, se aplican logaritmos para simplificar los cálculos y se escoge el parámetro que maximiza el valor de la función. Ese valor del parámetro será el estimador que se busca para el presente estudio.

4.2. Realización del Test Chi-Cuadrado: Bondad del ajuste

Para determinar el modelo de distribución que mejor se ajusta a las características de los datos de frecuencias y severidades observados se recurre al test Chi-Cuadrado. Este test es una prueba no paramétrica que permite determinar si una distribución empírica, como son los datos de frecuencias y siniestros registrados de cada clúster, proceden de una distribución teórica concreta bajo determinados parámetros que la caracterizan.

Por lo tanto, una vez que se han obtenido las estimaciones de los parámetros por Máxima Verosimilitud que caracterizan la distribución, el siguiente paso es comprobar la bondad del ajuste de los datos observados en la muestra a una distribución de probabilidad teórica concreta con los parámetros estimados.

Lo más importante a la hora de aplicar el Test Chi-Cuadrado es definir cuál es la hipótesis a contrastar. En este caso se definirá como hipótesis nula, H_0 , que los datos observados en el clúster,

tanto para frecuencia como para severidad, siguen la distribución teórica determinada con los parámetros ajustados previamente. Por el contrario, la hipótesis alternativa, H_1 , será el resultado de que el ajuste de la distribución teórica no sea adecuado para los datos observados.

Para desarrollar esta prueba no paramétrica, en primer lugar, se toman los datos de las variables y se procede a agruparlas de tal manera que no exista ningún intervalo que contenga menos de cinco observaciones. En caso de tener intervalos con menos de cinco frecuencias, deberá agruparse con otro intervalo hasta que no existan agrupaciones con menos de cinco registros. Esta condición se debe a que el Test ganará mayor precisión con estas agrupaciones y el resultado será más adecuado. Esto generará las frecuencias agrupadas bajo la condición mencionada, lo que se denomina los valores observados.

Una vez realizada la agrupación adecuada de los valores de la variable según sus frecuencias, se proyectan las probabilidades teóricas según la distribución de probabilidad que se desee ajustar con los parámetros que previamente se hayan estimado mediante el Método de Máxima Verosimilitud. Una vez calculadas las probabilidades teóricas según la distribución candidata de cada uno de los valores observados en las muestras, se multiplica cada una de esas probabilidades por el número total de observaciones que componen la muestra. Cada uno de los valores calculados son los valores teóricos.

Seguidamente, se procede a calcular el estadístico Chi-Cuadrado con la siguiente fórmula:

$$\chi_{\alpha, k-1}^2 = \sum_{i=1}^n \frac{(O_i - T_i)^2}{T_i}$$

[Fórmula 62]

Donde:

O_i son los valores observados.

T_i son los valores teóricos.

k son los grados de libertad del estudio. Se determina como el número de sucesos diferentes considerados después de realizar la agrupación de frecuencias bajo la condición anterior.

α es el nivel de significación al cual se realiza el estudio.

Finalmente, una vez obtenido el valor del estadístico Chi-Cuadrado con la fórmula anterior, se recurre a las tablas de la distribución Chi-Cuadrado para buscar el valor que toma el estadístico con el nivel de confianza al cual se realiza el estudio y los grados de libertad que resulten de las agrupaciones necesarias. Si el valor del estadístico que se ha calculado es inferior al valor en



tablas, entonces no se rechazará (es decir, podrá aceptarse) la hipótesis nula H_0 . En cambio, si el valor calculado es superior al valor que reflejan las tablas, el resultado del test de bondad del ajuste será que se rechaza que los datos del clúster sigan la distribución teórica bajo esos parámetros.

4.3. Comprobación del p-Valor

Una vez realizado el Test Chi-Cuadrado puede darse el caso de obtener más de una distribución de probabilidad que puede ajustarse al conjunto de datos de cada clúster. Por este motivo, se calcula el p-Valor para confirmar cuál será la distribución de probabilidad que mejor refleja el comportamiento de los datos y que resultará escogida para modelizar el comportamiento de ese clúster.

El p-Valor permitirá representar el resultado del contraste de hipótesis del experimento de una manera objetiva. En consecuencia, el valor de este parámetro refleja lo verosímil que es la muestra, es decir, la probabilidad de que se obtenga un valor del estadístico del contraste que sea al menos tan extremo como el observado, asumiendo que la hipótesis nula es verdadera. Otro enfoque del p-valor es que puede utilizarse para indicar cuánto contradice la muestra analizada la hipótesis alternativa. De este modo, los valores pequeños del p-Valor suponen que es poco probable obtener una muestra como la que se está analizando. En cambio, los valores altos indican que es frecuente volver a obtener una muestra como la analizada.

De esto cabe deducir que si el p-Valor es inferior al nivel de significación α , se rechazará la hipótesis nula y se aceptará la alternativa. En cambio, si el p-Valor es igual o superior al nivel de significación, no se rechazará la hipótesis nula.

Para el cálculo del p-Valor, se recurrirá a la función de distribución de la distribución Chi-Cuadrado y se buscará el valor del estadístico Chi-Cuadrado obtenido de los cálculos anteriores con sus correspondientes grados de libertad.



5. Distribución de reclamaciones totales mediante el método de Panjer

El último paso previo a la tarificación de seguros No Vida es la obtención de una distribución de pérdidas agregadas con sus correspondientes probabilidades, caracterizada por la distribución de probabilidad que siguen tanto las frecuencias como la severidad de los siniestros de los clústeres de la cartera de asegurados del producto estudiado.

Este epígrafe está dedicado en primer lugar a explicar la metodología para la discretización, realizada mediante el método *Rounding*, que permitirá convertir las distribuciones de probabilidad continuas asociadas a las severidades en una distribución discreta. Este paso es imprescindible para poder aplicar el algoritmo de Panjer, segundo punto del epígrafe, y que se realiza para obtener la distribución de pérdidas agregadas de cada clúster. Esta distribución de pérdidas agregadas será la utilizada para la tarificación final del producto en asegurador.

5.1. Discretización de las distribuciones continuas: Método *Rounding*

Una de las exigencias que tiene el método de Panjer para su aplicación es utilizar distribuciones de probabilidad discretas para obtener la función de probabilidad de las pérdidas agregadas. Esto supone realizar un proceso de discretización sobre las variables que siguen una distribución continua, como es el caso de las severidades en el presente estudio.

Con la finalidad de conseguir la discretización, el primer paso es definir una unidad de medida conveniente, h , denominada *span*. Esto permitirá asociar a cada tramo definido por los múltiplos de ese *span* las probabilidades discretas correspondientes. En este sentido, el usuario definirá en el aplicativo el *span* que desee aplicar para la discretización de la severidad, con un mínimo de 10 y un máximo de 5000, obteniendo un mayor grado de precisión cuanto menor sea el *span*, aunque implicará un tiempo de cálculo superior. El motivo de permitir que el usuario defina este parámetro otorgará mayor flexibilidad al aplicativo y podrá adaptarse el análisis de un producto asegurador según sus características.

El proceso de discretización óptimo debe permitir preservar los momentos representativos de la distribución continua, así como la forma de ésta. Existen diversas metodologías para alcanzar la discretización de distribuciones continuas, tales como el método *Rounding* o el Método *Local Moment Matching*.



En el presente TFM se ha escogido para la discretización de las distribuciones de severidad es el método *Rounding*. Si bien el método *Local Moment Matching* podría alcanzar un mayor grado de precisión, el tiempo computacional en el que se incurre debido a la complejidad de sus cálculos es muy superior. Además, los resultados alcanzados por *Local Moment Matching* suponen diferencias mínimas que apenas mejoran al método *Rounding*, lo cual justifica la elección de esta metodología para el proceso de discretización en el aplicativo desarrollado.

El método *Rounding* consiste en concentrar puntos de masa de probabilidad en los intervalos de discretización seleccionados por el usuario al especificar el *span*, de tal manera que se aplica la siguiente expresión¹ para conseguir la discretización de la función, siendo h el *span*:

$$f_0 = Pr\left(X < \frac{h}{2}\right) = F_x\left(\frac{h}{2} - 0\right)$$

$$f_j = Pr\left(jh - \frac{h}{2} \leq X < jh + \frac{h}{2}\right) = F_x\left(jh + \frac{h}{2} - 0\right) - F_x\left(jh - \frac{h}{2} - 0\right); \quad j = 1, 2, 3 \dots$$

[Fórmula 63]

Al realizar los cálculos se consigue redondear las cantidades asociadas a las severidades a los múltiplos de la unidad monetaria seleccionada que resulte más adecuada, definida por el usuario con el *span* de la distribución.

Aplicando la fórmula anterior se conseguirá la versión discretizada de la distribución de la variable severidad. Habiendo completado la modelización del comportamiento mediante las técnicas estadísticas explicadas anteriormente y con este punto completado, la siguiente fase consiste en implementar el método de Panjer para la obtención de la distribución de pérdidas agregadas de cada clúster.

5.2. Aplicación del algoritmo recursivo de Panjer

El algoritmo de Panjer para la construcción de la distribución de pérdidas agregadas es una alternativa a la metodología de cálculo por convolución. El cálculo mediante convoluciones implicaría un tiempo de cálculo desmesurado en comparación con este método, permitiendo alcanzar los mismos resultados de una manera más eficiente.

¹ La expresión matemática se obtiene del libro “*Loss Models; From data to decisions*” (4ª Edición, 2012), pp. 166-167.



El algoritmo desarrollado por Panjer puede aplicarse bajo determinadas condiciones, de las cuales cabe resaltar las siguientes:

1. Las distribuciones de frecuencias deben pertenecer a la clase (a,b,0) o a la clase (a,b,1).
2. La distribución de severidad debe ser discretizada previamente.

Ambas condiciones han sido tenidas en cuenta y analizadas a lo largo del presente documento.

Por tanto, el algoritmo de Panjer se define como un método recursivo que calcula la distribución exacta de un proceso compuesto, utilizando la distribución discreta de la clase (a,b,0) o (a,b,1) que mejor se ajuste a las frecuencias y la versión discretizada de la distribución de severidad.

La fórmula recursiva de Panjer puede diferenciarse según la clase a la que pertenezcan las distribuciones de probabilidad asociadas a las frecuencias. La fórmula general² aplicable para las distribuciones de la clase (a,b,1), siendo las variables frecuencia y siniestralidad independientes e igualmente distribuidas, es:

$$f_s(x) = \frac{[p_1 - (a + b)p_0] f_x(x) + \sum_{y=1}^{x \wedge m} \left(a + b \frac{y}{x}\right) f_x(y) f_s(x - y)}{1 - a f_x(0)} \quad \text{[Fórmula 64]}$$

Donde:

S es la cantidad total reclamada.

p_0 y p_1 son las probabilidades correspondientes a la distribución de la clase (a,b,1).

a y b son los parámetros que se obtienen de la reparametrización efectuada en el *epígrafe 3* según la distribución que mejor se ajuste a la variable frecuencia de siniestros.

$f_x(x)$ es la probabilidad discreta asociada a la severidad.

$x \wedge m$ representa el mínimo entre x y m . La variable x es la dimensión del vector que representa los valores discretizados de la severidad y m es la cuantía más elevada expresada en unidades monetarias que puede ser reclamada.

En el caso de que la distribución que mejor se ajuste pertenezca a la clase (a,b,0), la fórmula a aplicar se simplificaría, quedando la expresión siguiente:

$$f_s(x) = \frac{\sum_{y=1}^{x \wedge m} \left(a + b \frac{y}{x}\right) f_x(y) f_s(x - y)}{1 - a f_x(0)} \quad \text{[Fórmula 65]}$$

² La expresión matemática se obtiene del libro “Loss Models; From data to decisions” (4º Edición, 2012), pp. 161-166. Del mismo libro se obtiene la *Fórmula 65*.



Para utilizar este algoritmo, en primer lugar hay que calcular cuál será el valor de la función en 0, es decir, $f_s(0)$. En la *Tabla 3* se resume la forma de calcular este valor para cada una de las distribuciones de frecuencias estudiadas.

Tabla 3. Cálculo del valor de $f_s(0)$ según la distribución de probabilidad de las frecuencias.

Distribuciones de probabilidad	Cálculo de $f_s(0)$
Binomial	$[1 + p(f_x(0) - 1)]^n$
Binomial ZM y ZT	$p_0^M + (1 - p_0^M) \frac{[1 + p(f_x(0) - 1)]^n - (1 - p)^n}{1 - (1 - p)^n}$
Binomial Negativa	$[1 + \beta(1 - f_x(0))]^{-r}$
Binomial Negativa ZM y ZT	$p_0^M + (1 - p_0^M) \frac{[1 + \beta(1 - f_x(0))]^{-r} - (1 + \beta)^{-r}}{1 - (1 + \beta)^{-r}}$
Geométrica	$[1 + \beta(1 - f_x(0))]^{-1}$
Geométrica ZM y ZT	$p_0^M + (1 - p_0^M) \frac{f_x(0)}{1 + \beta(1 - f_x(0))}$
Poisson	$e^{[\lambda(f_x(0)-1)]}$
Poisson ZM y ZT	$p_0^M + (1 - p_0^M) \frac{e^{\lambda f_x(0)} - 1}{e^\lambda - 1}$

Fuente: *Elaboración propia a partir de la información obtenida del libro “Loss Models: From data to decisions” (2012).*

Aplicando la metodología propuesta por Panjer, se obtendrán la distribución de pérdidas agregadas total y los estadísticos representativos de cada clúster, definiendo tanto la media como varianza y desviación típica. Estos datos orientativos dan una primera aproximación al usuario de la esperanza de los costes y cómo de representativa es esa media obtenida. En consecuencia, podría adoptarse una política de tarificación que se ajuste al comportamiento de cada clúster y



además permita tener una primera aproximación de la suficiencia de las cuantías cobradas en comparación con la simulación de severidad que se propondrá más adelante.

Una vez definido el procedimiento para obtener la distribución de pérdidas agregadas totales y habiendo comprendido el comportamiento de cada uno de los clústeres, se procede a aplicar los cálculos para la tarificación de seguros.

6. Tarificación de seguros No Vida

Todos los pasos que han sido explicados anteriormente forman parte del proceso de tarificación que una entidad aseguradora debería seguir para aplicar el algoritmo de Panjer. Sin ellos no podrían haberse obtenido los dos elementos principales que formarán parte de la expresión general de tarificación implementada en el aplicativo: esperanza de la severidad y medida de riesgo.

Mediante la aplicación de conocimientos estadísticos se ha modelizado el comportamiento que siguen los asegurados pertenecientes a cada clúster en función de la distribución de probabilidad que mejor representaba su comportamiento. En el punto tratado en el presente epígrafe, entra en juego los diversos enfoques actuariales para la tarificación de seguros y la gestión de riesgos inherentes a la actividad aseguradora.

Este último punto abre camino para hablar de perfiles de riesgo, políticas asumidas y gestión de éstos por parte de la entidad. En el presente TFM se ha incluido un epígrafe adicional que permite la simulación de severidades para hacer frente a las potenciales reclamaciones que los asegurados pueden realizar en un periodo temporal, atendiendo al comportamiento mostrado por la distribución de pérdidas calculada.

Volviendo a lo referente al apartado que ocupa este epígrafe, la tarificación consiste en determinar una cuantía monetaria fija por la cual la entidad aseguradora está dispuesta a asumir el riesgo que el asegurado desea transferir para prevenir una posible pérdida financiera futura, el cual estará sujeta a aleatoriedad. Con la acumulación del total de primas recibidas la entidad deberá hacer frente a los siniestros que se presenten y que estén dentro de las condiciones generales y particulares recogidas en la póliza.

A continuación, y con motivo de introducir un nuevo concepto que se utilizará posteriormente en la tarificación, se define la idea de medida de riesgo y, más concretamente, el concepto de *Value-at-Risk*. Esta será la medida de riesgo escogida para el aplicativo desarrollado en el presente TFM para la tarificación de seguros.



6.1. *Value-at-Risk*: medida de riesgo para la tarificación de seguros No Vida

La medida de riesgo es el concepto que hace referencia a las potenciales pérdidas económicas que una entidad puede sufrir en un escenario dado, es decir, refleja la cantidad total, expresada en unidades monetarias, que una compañía está en riesgo de perder en un determinado periodo de tiempo como consecuencia de un escenario cuya probabilidad de materializarse se entiende razonable. Se identificará en el presente TFM la medida de riesgo por medio de $\rho(-y)$.

Existen diversas formas de representar este riesgo con varias medidas como el *Value-at-Risk* (*VaR*), el *Conditional Value-at-Risk*, el *Conditional Value-at-Risk ponderado*, o mediante la aplicación de medidas de distorsión como la *Dual Power Transform* o la medida de *Wang*. En el presente TFM se ha fijado como medida de riesgo el *VaR*, dado que es una medida muy extendida en la práctica actuarial y financiera, estando presente tanto en la directiva de regulación de Basilea, en el caso de entidades financieras, y de Solvencia para las compañías aseguradoras.

El *VaR* o valor en riesgo es una medida de riesgo que se define como el percentil que refleja la mayor pérdida que una entidad podría experimentar a un determinado nivel de confianza previamente definido.

Siguiendo la definición anterior, hay tres elementos que la componen: el tamaño total de la pérdida, el nivel de confianza y el periodo de tiempo. Definiendo dos de estas variables, puede obtenerse la tercera. El objetivo en el presente TFM será medir la máxima pérdida esperada posible (tamaño de la pérdida) que puede darse en la cartera del producto asegurador analizado para un determinado nivel de confianza que se especificará en el aplicativo y asumiendo como horizonte temporal aquel que se estipule en el contrato de seguro.

Desde el punto de vista de gestión de una entidad aseguradora, el *VaR* es la cantidad de capital que los accionistas invertirían con el objetivo de limitar la probabilidad de bancarrota en caso de que las condiciones fueran adversas. En cualquier caso, no puede interpretarse esta medida como cuánta es la pérdida que una entidad puede experimentar, sino que con un determinado nivel de confianza las pérdidas no superarán la cifra facilitada por el *VaR* en el periodo de tiempo asumido.

La forma de calcular esta medida de riesgo es recurriendo a la función de distribución de pérdidas agregadas calculada por el algoritmo de Panjer. Disponiendo de esa distribución, es sencillo recurrir al percentil especificado por el usuario del aplicativo y que refleja la cuantía máxima esperada al nivel de confianza al cual se efectúa el estudio.



6.2. Principio general para la tarificación de seguros y sus extensiones

Toda vez que se ha caracterizado y evaluado el comportamiento de los miembros presentes en cada clúster y los riesgos que entrañan, se puede proceder a ajustar una tarifa que permita la transferencia de riesgos y flujos monetarios entre las partes.

Para efectuar la tarificación en el presente se ha recurrido a la siguiente expresión general de la cual puede derivarse otras formas de tarificación:

$$\Pi(y) = E(y)(1 + k_1) + k_2\rho(-y) \quad \text{[Fórmula 66]}$$

Donde:

$\Pi(y)$ representa la cuantía a cobrar a cada asegurado, la prima.

$E(y)$ representa la esperanza del coste de cada asegurado perteneciente al clúster, obtenida de la distribución de pérdidas agregadas de Panjer.

k_1 es el recargo de seguridad aplicado directamente sobre la esperanza de la severidad. Bajo determinados enfoques, este valor puede llegar a ser -1.

$\rho(-y)$ es una medida de riesgo escogida para tarificar.

k_2 es el recargo de riesgo aplicado directamente sobre la medida de riesgo seleccionada. Este parámetro toma valores superiores o iguales a 0.

De la expresión general anterior para la tarificación de seguros pueden derivarse dos principios adicionales de tarificación de seguros No Vida.

Por un lado, en caso de fijar el parámetro del recargo de seguridad con un valor superior a 0 y asumiendo que el parámetro de recargo por riesgo sea nulo, se obtiene el Principio de Prima de Valor Esperado. Suele utilizarse en la práctica actuarial, pero tiene como inconveniente que sólo utiliza la esperanza de la distribución para obtener la tarifa del seguro en lugar de utilizar la distribución completa del coste agregado.

La expresión matemática del Principio de Prima de Valor Esperado será:

$$\Pi(y) = E(y)(1 + k_1) \quad \text{[Fórmula 67]}$$



Por otro lado, si el parámetro del recargo de seguridad toma como valor -1 y el parámetro del recargo de riesgo es superior a 0, surge el Principio de Prima de la Función del Riesgo. En este caso, la esperanza no será considerada como un elemento de tarificación y se recurrirá a alguna medida de riesgo que sea representativa de los riesgos soportados por la entidad aseguradora.

Por lo tanto, la expresión del modelo de tarificación bajo este principio es:

$$\Pi(y) = k_2 \rho(-y)$$

[Fórmula 68]

Mediante cualquiera de los principios anteriormente citados puede obtenerse una tarifa que recoja los componentes que caracterizan el comportamiento del asegurado. Queda a completa disposición del usuario del aplicativo según las políticas de riesgos y comerciales de la entidad la fijación de los recargos de seguridad y riesgo que considere.

Finalmente, el usuario puede plantearse la pregunta ¿son las primas cobradas a los clientes suficientes para hacer frente al total de reclamaciones que puedan presentarse? Por este motivo, y para tener una visión integral de lo adecuadas que son las tarifas fijadas, se ha incluido un punto adicional para la simulación de diversos escenarios que permitan evaluar la adecuación de las tarifas diseñadas por la entidad.



7. Simulación de severidad: Adecuación de las primas cobradas

Las entidades aseguradoras deben dotar una cuantía que sea suficiente para hacer frente a las reclamaciones potenciales que los clientes exigirán por los siniestros acaecidos durante el periodo de vigencia del contrato firmado. Este es el concepto de reserva, que deberá cubrir tanto las cuantías estimadas que serán reclamadas como los gastos implícitos para satisfacer las obligaciones contraídas. Este cálculo es esencial dentro de la actividad aseguradora y lo adecuado que sea supone estar en condiciones de poder cumplir las obligaciones adquiridas. Es parte de los cálculos que desarrolla el departamento actuarial.

En el aplicativo desarrollado en el presente TFM, se ha optado por utilizar la función de distribución de las pérdidas agregadas por el método de Panjer. En base a esta distribución se simulan escenarios aleatorios que dan como resultado una distribución de severidades agregadas.

El método utilizado consiste en, una vez construida la función de distribución de las pérdidas agregadas, simular números aleatorios entre 0 y 1 interpretando éstos como probabilidades. Cada número aleatorio simulado se compara con la probabilidad acumulada en la función de distribución de pérdidas agregadas obtenida por Panjer, la cual estará asociada a una cuantía de severidad que será la que se asuma como la reclamación de ese individuo. Este proceso se realiza para cada uno de los asegurados que pertenezcan al clúster, lo cual ofrecerá una simulación de un escenario del comportamiento de cada uno de los asegurados y la reclamación que efectuaría en ese periodo de tiempo. Una vez se ha realizado la simulación de cada uno de los asegurados que conforma el clúster, se suman todas las cuantías individuales simuladas, lo cual dará una estimación de cuál será la reclamación total de ese clúster.

Con el objetivo de tener una simulación consistente, el aplicativo está diseñado para efectuar este procedimiento tantas veces como el usuario desee. Obviamente, cuanto mayor sea el número de simulaciones, más fiable será el dato al tener una muestra más amplia sobre la siniestralidad que puede presentarse en varios escenarios. De este modo podrá valorarse de manera más fiable la capacidad de solvencia de la empresa. Finalmente, podrá construirse una distribución y podrá calcularse un determinado percentil de ésta y obtener el valor de la reclamación total simulada que corresponda a ese percentil de la distribución.



Con este resultado, se tiene una primera aproximación de las reclamaciones totales y además es posible contrastar la capacidad que tiene para hacer frente con las primas totales cobradas a las posibles reclamaciones que puedan surgir. Esto se realiza obteniendo el cociente entre las primas totales cobradas y el valor de las pérdidas simuladas para ese percentil escogido.

El departamento encargado de la modelización de las reservas puede utilizar la simulación realizada en el aplicativo para hacer estimaciones de la cantidad que la empresa deberá dotar para hacer frente a las reclamaciones que se presentarán durante la vigencia de los contratos.

Para la constitución de las reservas existen distintas técnicas, como *Chain Ladder*, para proyectar el patrón de siniestros reclamados a la entidad en relación al ramo al que pertenezca el producto. No obstante, es necesario tener un histórico de reclamaciones del producto para poder efectuar una proyección que refleje realmente el comportamiento de reclamaciones de la cartera.

En caso de no disponer de un histórico suficientemente amplio, la entidad puede realizar estudios para obtener un *benchmark* de mercado y utilizarlo para aproximar una cuantía de reservas acordes al producto. A medida que se obtenga mayor conocimiento sobre el comportamiento de la cartera, podrá ajustarse de mejor manera el comportamiento y se tendrá una base suficiente para poder utilizar el histórico del producto.



8. Resultados alcanzados: Ejemplo ilustrativo

A continuación, se ilustra un ejemplo del funcionamiento del aplicativo diseñado y sobre el cual se aplican todas las explicaciones teóricas anteriores.

8.1. Inputs del aplicativo

El aplicativo requiere, en primer lugar, la introducción de una base de datos almacenada en Excel. Esta base de datos deberá contener en la primera columna de la hoja las frecuencias de siniestros. En la segunda columna deberán figurar las reclamaciones que han surgido durante el periodo analizado. Finalmente, se introduce una tercera columna en la cual se identificará el clúster al cual pertenecen los asegurados con un número identificativo (1, 2, 3...). Una vez que la base de datos ha sido seleccionada, aparecerá el nombre en el recuadro del aplicativo destinado a tal fin.

A continuación, es necesario introducir los parámetros para realizar el análisis de los clústeres y la tarificación. Los inputs para el análisis de los asegurados que componen la base de datos a estudiar serán:

- **Número de clústeres:** Corresponde al número de clústeres que la empresa ha definido para analizar el producto. En el caso del ejemplo simulado, la empresa define un total de 3 clústeres para el producto asegurador.
- **Nivel de significación:** Nivel de referencia α al cual se realiza el Test Chi-Cuadrado. En el ejemplo mostrado a continuación se tomará un valor típico para estudios del 5%.
- **Modificación de la probabilidad en o:** Para el caso de las distribuciones de probabilidad de la clase (a,b,1), se aplica una modificación de la probabilidad en el punto o de la distribución para efectuar el ajuste en base a esta modificación. Se ha optado por fijar este valor para el ejemplo en 0,35.

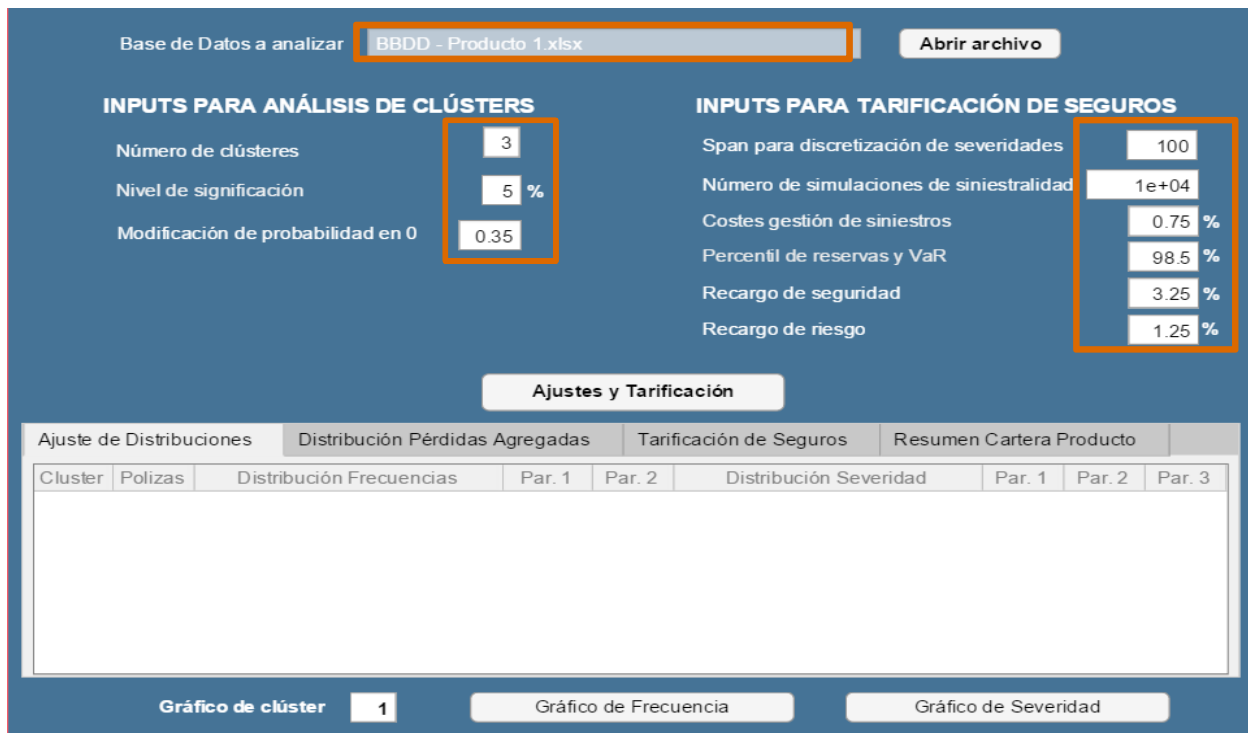
Por otro lado, los inputs para la tarificación de seguros serán:

- **Span para la discretización de severidades:** Consiste en determinar una unidad monetaria que el usuario considere adecuada con el fin de discretizar la distribución de pérdidas. Se calcularán múltiplos de esa unidad a los que se asignará la probabilidad correspondiente. En el caso expuesto, se fijará un *span* de 100 unidades monetarias.

- **Número de simulaciones de siniestralidad.** Hace referencia al número de veces que se simulará las siniestralidades utilizando el método descrito en el apartado anterior. Para el caso propuesto, se efectuarán 10.000 simulaciones en cada clúster.
- **Costes de gestión de siniestros.** Cada vez que un siniestro se presenta, acarrea una serie de gastos administrativos y de gestión asociados a la tramitación del expediente. En este caso, se fija como gastos de gestión un 0,75% de la severidad.
- **Percentil de reservas y VaR.** Para ambos valores se utilizará el mismo percentil. En el ejemplo se utilizará un percentil del 98,5%.
- **Recargo de seguridad.** Será el porcentaje que se recarga adicionalmente sobre la esperanza del coste (k_1). Se utilizará un 3,25% para el presente ejercicio.
- **Recargo de riesgo.** Porcentaje (k_2) aplicado directamente sobre la medida de riesgo VaR. Se opta por fijarlo en un 1,25%.

La *Imagen 1* que figura a continuación resume la información necesaria a introducir a modo de ejemplo.

Imagen 1. Introducción de inputs para análisis de clústeres y tarificación de seguros.



Base de Datos a analizar:

INPUTS PARA ANÁLISIS DE CLÚSTERS

Número de clústeres:

Nivel de significación: %

Modificación de probabilidad en 0:

INPUTS PARA TARIFICACIÓN DE SEGUROS

Span para discretización de severidades:

Número de simulaciones de siniestralidad:

Costes gestión de siniestros: %

Percentil de reservas y VaR: %

Recargo de seguridad: %

Recargo de riesgo: %

Ajuste de Distribuciones | **Distribución Pérdidas Agregadas** | Tarificación de Seguros | Resumen Cartera Producto

Cluster	Polizas	Distribución Frecuencias	Par. 1	Par. 2	Distribución Severidad	Par. 1	Par. 2	Par. 3

Gráfico de clúster

8.2. Resultados del análisis y tarificación

Una vez que la base de datos y los inputs para el análisis y tarificación han sido definidos, se pulsa el botón “Ajustes y Tarificación”. En ese momento, se ejecutará el código desarrollado y la consiguiente realización de los cálculos que se han presentado teóricamente a lo largo del documento.

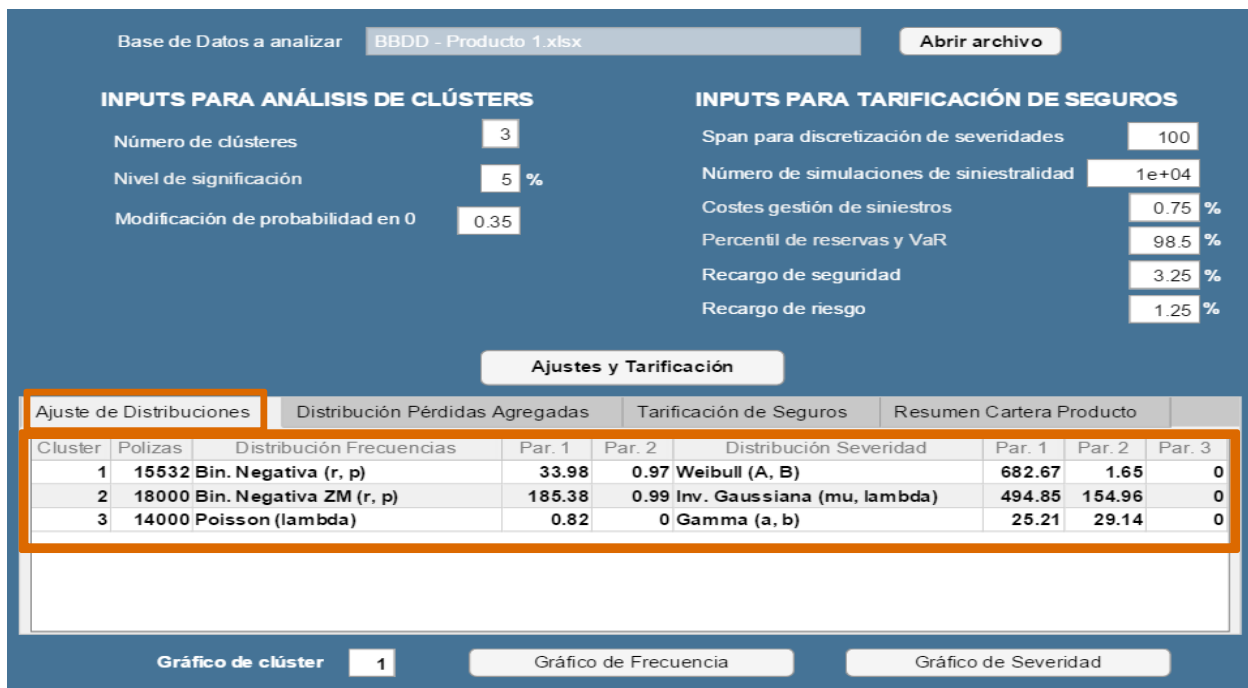
Los resultados alcanzados por el aplicativo se muestran a lo largo de cuatro pestañas, las cuales se explican a continuación.

8.2.1. Resultados de ajustes a distribución de probabilidad teóricas

En primer lugar, se procede al ajuste de las distribuciones de probabilidad y se determina aquella que mejor representa el comportamiento de los asegurados, tanto para la frecuencia como para la severidad.

Los resultados se muestran en la *Imagen 2*, reflejando el número de pólizas identificadas en cada clúster, las distribuciones discretas asociadas a las frecuencias de siniestros y las distribuciones continuas que mejor se ajustan a las severidades de cada clúster. También se especifican los parámetros hallados por Máxima Verosimilitud para cada distribución.

Imagen 2. Ajuste de distribuciones de frecuencias y severidades de cada clúster.



The screenshot displays the 'Ajustes y Tarificación' interface. At the top, it shows the data source 'BBDD - Producto 1.xlsx' and an 'Abrir archivo' button. Below are two sections of input parameters:

- INPUTS PARA ANÁLISIS DE CLÚSTERS:**
 - Número de clústeres: 3
 - Nivel de significación: 5 %
 - Modificación de probabilidad en 0: 0.35
- INPUTS PARA TARIFICACIÓN DE SEGUROS:**
 - Span para discretización de severidades: 100
 - Número de simulaciones de siniestralidad: 1e+04
 - Costes gestión de siniestros: 0.75 %
 - Percentil de reservas y VaR: 98.5 %
 - Recargo de seguridad: 3.25 %
 - Recargo de riesgo: 1.25 %

A central button labeled 'Ajustes y Tarificación' is visible. Below it, a tabbed interface shows four tabs: 'Ajuste de Distribuciones' (selected), 'Distribución Pérdidas Agregadas', 'Tarificación de Seguros', and 'Resumen Cartera Producto'. The 'Ajuste de Distribuciones' tab contains the following table:

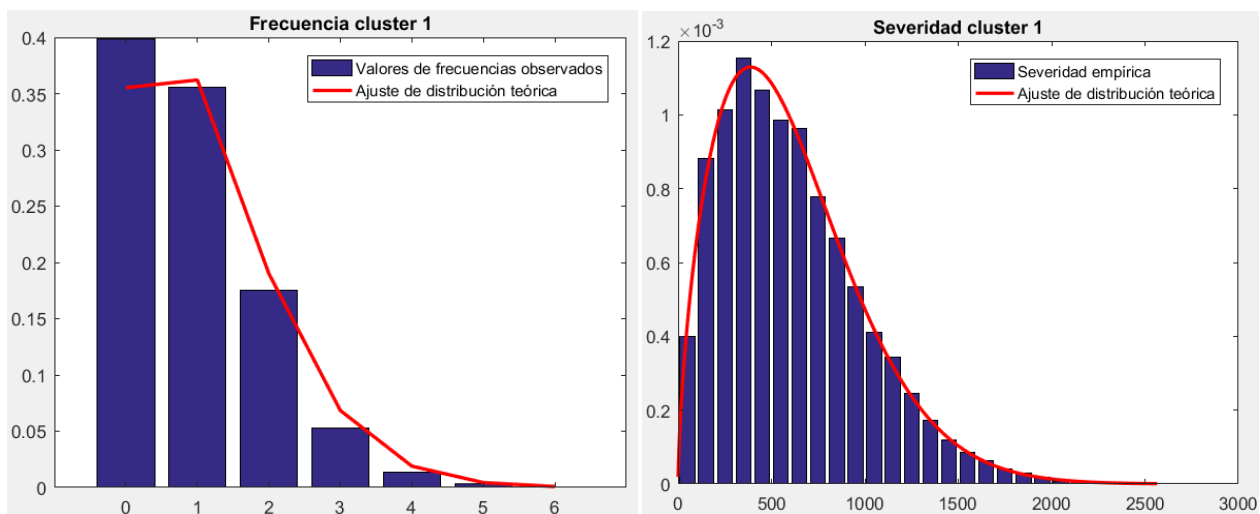
Clúster	Polizas	Distribución Frecuencias	Par. 1	Par. 2	Distribución Severidad	Par. 1	Par. 2	Par. 3
1	15532	Bin. Negativa (r, p)	33.98	0.97	Weibull (A, B)	682.67	1.65	0
2	18000	Bin. Negativa ZM (r, p)	185.38	0.99	Inv. Gaussiana (mu, lambda)	494.85	154.96	0
3	14000	Poisson (lambda)	0.82	0	Gamma (a, b)	25.21	29.14	0

At the bottom, there are three buttons: 'Gráfico de clúster' (with a dropdown set to '1'), 'Gráfico de Frecuencia', and 'Gráfico de Severidad'.

Además de los resultados del ajuste de manera numérica, se muestra a continuación un ejemplo de cómo resulta el ajuste de manera visual, proyectando el gráfico de las frecuencias y severidades observadas de cada clúster y el ajuste a la distribución de probabilidad teórica. Estos gráficos se obtienen pulsando sobre los botones “Gráfico de Frecuencia” y “Gráfico de Severidad”.

Para ilustrar un ejemplo, se muestra el comportamiento de las frecuencias observadas del Clúster 1, cuya distribución teórica es una Binomial Negativa de parámetro r de 33,98 y p con valor 0,97. En el caso de la distribución de severidad, sigue una distribución teórica de Weibull de parámetro A de 682,67 y B de 1,65.

Imagen 3. Ajuste de distribuciones a valores observados de cada clúster.

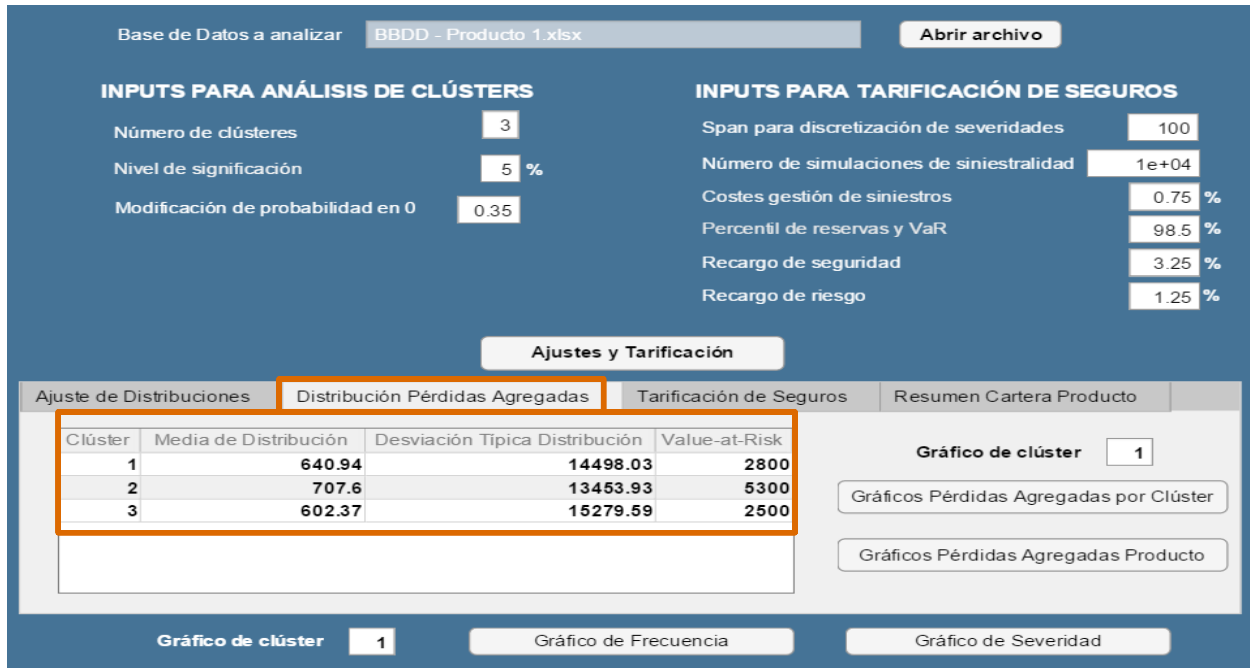


8.2.2. Obtención de la distribución de pérdidas agregas por el algoritmo de Panjer

La siguiente pestaña del aplicativo ofrece la información resultante de aplicar el algoritmo de Panjer para calcular las distribuciones de pérdidas agregadas. Por un lado, muestra cuáles son los estadísticos media y desviación típica de la distribución de pérdidas calculada. Además, utilizando la propia función de distribución calculada por el algoritmo de Panjer, puede obtenerse el VaR , que posteriormente se utilizará para la tarificación de cada clúster.

En la *Imagen 4* se muestran los resultados alcanzados en esta simulación.

Imagen 4. Resultados de cálculo de la distribución de pérdidas agregadas mediante método de Panjer.



De este resultado mostrado por el aplicativo puede observarse como el Clúster 2 presenta un comportamiento más arriesgado, puesto que el *VaR* es muy superior con respecto a los otros dos clústeres. En otras palabras, es el grupo homogéneo que tendería, a priori, a presentar unas reclamaciones más elevadas. Sin embargo, es precisamente ese clúster el más homogéneo en su comportamiento, ya que la desviación típica es menor que en los otros dos.

Puede comprobarse que mediante la aplicación del algoritmo propuesto por Panjer se obtiene un resultado prácticamente idéntico al que se conseguiría utilizando como esperanza de pérdidas agregadas el producto de las esperanzas de las variables frecuencia y severidad.

$$E(S) = E(N)E(X)$$

[Fórmula 69]

Siguiendo el ejemplo del Clúster 1, se procede a realizar el cálculo del producto de esperanzas de ambas variables para obtener la esperanza del coste agregado.

$E(N)$ hace referencia a la esperanza de la variable frecuencia de siniestros. En el caso del Clúster 1, la distribución sigue una distribución Binomial Negativa cuya r es 33,98 y p es 0,97. Se recurre a la *Fórmula 9* para calcular la esperanza de la media, tal como figura a continuación.

$$E(N) = \frac{33,98 (1 - 0,97)}{0,97} = 1,0509$$

Por otro lado, $E(X)$ representa la esperanza de la variable severidad. Se calcula recurriendo a la *Fórmula 57* que expresa el cálculo de la esperanza de una distribución Weibull, en este caso de parámetro A 682,67 y B 1,65. El resultado de la esperanza de la severidad es:

$$E(X) = 682,67 \Gamma(1 + 1,65^{-1}) = 610,45$$

Finalmente, recurriendo a la *Fórmula 69* se obtiene la expresión de la esperanza del coste agregado:

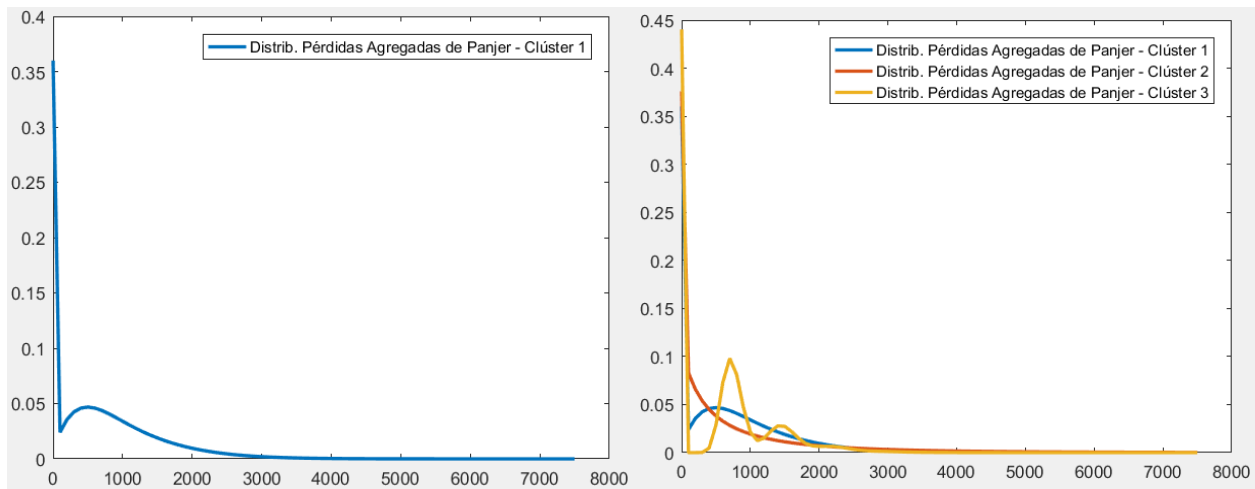
$$E(S) = 1,0509 \times 610,45 = 641,52$$

La diferencia entre realizar el cálculo por Panjer y aplicando el producto de medias es de 0,58 unidades monetarias. Esta pequeña diferencia puede deberse al *span* utilizado para efectuar la discretización, dado que a medida que se incrementa el *span* tiende a perderse algo de precisión en el cálculo.

De igual modo que en la pestaña anterior, puede obtenerse la distribución de pérdidas agregadas de una manera más visual generando los gráficos de cada clúster y de la cartera. Para ello, habrá que pulsar en el botón “*Gráficos de Pérdidas Agregadas por Clúster*” y “*Gráfico Pérdidas Agregadas Producto*”, respectivamente.

En la *Imagen 5* se muestra un ejemplo tomando como referencia de nuevo el Clúster 1 para el caso individual y el gráfico correspondiente a la distribución de pérdidas agregadas del producto.

Imagen 5. Distribución de pérdidas agregadas por clúster (izquierda) y del producto asegurador (derecha).



8.2.3. Tarificación de seguros para cada clúster

Seguidamente, se muestran los resultados de tarificación del producto según el clúster al que pertenece aplicando los recargos previamente especificados en el aplicativo.

Para poder obtener una perspectiva sobre la adecuación de las primas, se ha optado por mostrar los cálculos relativos al volumen total al que asciende las primas cobradas y la siniestralidad simulada para cada clúster. Puede obtenerse un ratio de estas dos magnitudes para comprobar cómo de convenientes son las primas cobradas en relación a la severidad que puede registrarse a un nivel de confianza del 98,5%.

A continuación, en la *Imagen 6* pueden comprobarse los resultados alcanzados.

Imagen 6. Resultados de la tarificación de cada clúster.



The screenshot shows a software interface with the following sections:

- Base de Datos a analizar:** BBDD - Producto 1.xlsx (with 'Abrir archivo' button)
- INPUTS PARA ANÁLISIS DE CLÚSTERS:**
 - Número de clústeres: 3
 - Nivel de significación: 5 %
 - Modificación de probabilidad en 0: 0.35
- INPUTS PARA TARIFICACIÓN DE SEGUROS:**
 - Span para discretización de severidades: 100
 - Número de simulaciones de siniestralidad: 1e+04
 - Costes gestión de siniestros: 0.75 %
 - Percentil de reservas y VaR: 98.5 %
 - Recargo de seguridad: 3.25 %
 - Recargo de riesgo: 1.25 %
- Ajustes y Tarificación:**
 - Ajuste de Distribuciones
 - Distribución Pérdidas Agregadas
 - Tarificación de Seguros** (highlighted)
 - Resumen Cartera Producto

The 'Tarificación de Seguros' table is as follows:

Clúster	Pólizas	Esperanza siniestro	Prima cobrada	Total Primas cobradas	Siniestralidad simulada	Ratio Primas/siniestralidad
1	15532	640.94	696.77	10822231.64	10754558.75	1.0063
2	18000	707.6	796.85	14343300	14083187.63	1.0185
3	14000	602.37	653.19	9144660	8694170.88	1.0518

At the bottom, there are buttons for 'Gráfico de clúster' (set to 1), 'Gráfico de Frecuencia', and 'Gráfico de Severidad'.

Puede apreciarse que en los tres clústeres en los que la entidad ha dividido la cartera se ha obtenido un ratio superior a 1, lo que se interpreta como que las primas cobradas son suficientes para hacer frente al escenario simulado de un *VaR* al 98,5%, incluyendo también los costes de gestión de los siniestros. Esto permite concluir que, a priori y bajo las condiciones de simulación y tarificación empleadas, el precio cobrado por el seguro sería adecuado y que el incremento o decremento de los recargos aplicados dependerá de las estrategias comerciales y la gestión de riesgos que la entidad defina.

8.2.4. Resumen de la cartera del producto asegurador

Finalmente, el usuario dispone en el último apartado del aplicativo un resumen del producto en global para evaluar la viabilidad económica del mismo, así como las cuantías implícitas en la emisión de dichos contratos.

En él, puede comprobarse el total de pólizas del producto, la esperanza de la siniestralidad total, el total de primas cobradas, el ratio de las primas cobradas sobre la siniestralidad esperada (que debería ser superior a 1 para garantizar que se está cobrando una prima superior a la esperanza de la severidad), la siniestralidad total simulada y la evaluación del total de primas cobradas sobre la severidad total simulada del producto.

La *Imagen 7* muestra a continuación la evaluación final del producto.

Imagen 7. Resumen de la tarificación global del riesgo



Base de Datos a analizar: BBDD - Producto 1.xlsx Abrir archivo

INPUTS PARA ANÁLISIS DE CLÚSTERES

Número de clústeres: 3

Nivel de significación: 5 %

Modificación de probabilidad en 0: 0.35

INPUTS PARA TARIFICACIÓN DE SEGUROS

Span para discretización de severidades: 100

Número de simulaciones de siniestralidad: 1e+04

Costes gestión de siniestros: 0.75 %

Percentil de reservas y VaR: 98.5 %

Recargo de seguridad: 3.25 %

Recargo de riesgo: 1.25 %

Ajustes y Tarificación

Ajuste de Distribuciones |
 Distribución Pérdidas Agregadas |
 Tarificación de Seguros |
 Resumen Cartera Producto

A continuación, se muestran los resultados agregados del proceso de tarificación de la cartera de pólizas para el producto seleccionado.

Total Pólizas	Siniest Total Esperada	Total Primas cobradas	Primas/Sin. Esperada	Siniestralidad Total simulada	Primas/Sin. Simul.
47532	31125015.12	34310191.64	1.1023	33531917.26	1.0232

Gráfico de clúster: 1 |
 Gráfico de Frecuencia |
 Gráfico de Severidad

En conclusión, bajo las condiciones de tarificación y simulación especificadas al inicio, las primas totales cobradas por este producto garantizarían hacer frente a las severidades que puedan presentarse durante el periodo de vigencia del contrato. Concretamente, el ratio de primas cobradas sería de 1,0232, el cual quedaría sujeto a evaluación por parte del usuario para opinar de la adecuación de las primas según las políticas comerciales y estratégicas.



9. Conclusiones finales

El presente Trabajo de Fin de Máster ha tratado de presentar un enfoque avanzado que permita la tarificación de seguros del ramo No Vida aplicando metodología de inferencia estadística, el algoritmo recursivo de Panjer y enfoques actuariales bajo el criterio y opinión experta del usuario del aplicativo.

En base al ejercicio simulado anteriormente, puede comprobarse cómo el aplicativo define, en primera instancia, cuál es el modelo de probabilidad que mejor se adapta a cada clúster, recurriendo en el caso de las frecuencias de siniestros a distribuciones de la clase $(a,b,0)$ y $(a,b,1)$. Se demuestra así que la implantación de la clase $(a,b,1)$ tiene gran importancia a la hora de estudiar el comportamiento de los asegurados, puesto que hay casos en los que las distribuciones clásicas en el mundo actuarial pertenecientes a la clase $(a,b,0)$ pueden no ser del todo correctas para comprender el comportamiento de los asegurados. Quedarían cumplidos así los dos primeros objetivos planteados inicialmente en la introducción del presente documento.

Seguidamente, se ha procedido a la implementación, previa discretización de las distribuciones continuas de severidad, del algoritmo recursivo de Panjer para aunar en una única distribución las frecuencias y las severidades de cada clúster. Esto permite disponer de información muy valiosa para la tarificación a través de los estadísticos de ésta. Pero el alcance de la distribución de Panjer no se limita únicamente a la tarificación de seguros, si no que en el presente aplicativo se amplía para obtener el valor en riesgo (VaR) de cada clúster y la obtención de una función de distribución de pérdidas agregadas para analizar la severidad potencial que pueda darse durante la vigencia de los contratos. Efectuando este punto, se completa el tercer objetivo planteado.

Como continuación de lo anterior, se procede a simular números aleatorios para todos los clientes tantas veces como el usuario considere oportuno para obtener una evaluación consistente de la adecuación de las primas cobradas. Una vez finalizada la simulación, se comprueba mediante el ratio de primas cobradas y severidad simulada cómo de adecuada es la tarifa frente a una simulación de pérdidas a un determinado nivel de confianza, alcanzando así el cuarto objetivo inicialmente planteado.

Una vez conocida la esperanza de los siniestros agregados y el VaR gracias al algoritmo de Panjer, se aplica el modelo de tarificación que el usuario considere más acorde según su juicio, aplicando los recargos que satisfagan las condiciones exigidas a nivel interno según las estrategias



de la entidad aseguradora, y externo, en función de las exigencias de los reguladores para garantizar la viabilidad económica de la empresa y poder hacer frente a los compromisos adquiridos con terceras partes. En este punto se alcanza el quinto objetivo propuesto.

Los dos últimos apartados del aplicativo muestran la información de lo adecuadas que resultan las tarifas en relación a las simulaciones efectuadas y dan un primer paso para que el usuario evalúe la conveniencia o no de vender el seguro o replantearlo con un nuevo enfoque, ya sea a nivel técnico o comercial. Así, quedarían satisfechos los puntos sexto y séptimo de los objetivos.

Como conclusión, el presente TFM ha cumplido todos los objetivos que se plantearon al inicio y queda demostrada la importancia de los análisis estadísticos y la implantación de enfoques alternativos frente a modelos más clásicos, como se entiende el método de Panjer frente a las convoluciones, en la práctica actuarial.

Por último, apuntando a nuevas vías de investigación, recientes estudios demuestran que existe cierta correlación entre las variables frecuencia y severidad, es decir, no podría asumirse completa independencia entre ambas. En este sentido, el presente trabajo supone una primera pieza para la tarificación de seguros y podría ser interesante ampliar en el futuro un análisis de las correlaciones mostradas entre ambas variables aleatorias.

Bibliografía

Libros

Klugman, S. A.; Panjer, H. H.; Willmot, G.E.; (2012); “*Loss Models: From Data to Decisions*” (4ª Edición); John Wiley & Sons, INC.; ISBN 978-1 -118-31532-3.

Tse, Y.-K.; (2009); “*Nonlife Actuarial Models: Theory, Methods and Evaluation*” (1ª Edición); Cambridge University Press; ISBN-13 978-0-521-76465-0.

Artículos y publicaciones

Revista Varianza nº9; (2012); Mancilla Flores, E. M.; “*Algoritmo de Panjer*” [Documento disponible en línea]. Recurso electrónico disponible en el siguiente enlace: http://www.revistasbolivianas.org.bo/scielo.php?pid=S9876-67892012000100007&script=sci_arttext [Visitado el 3 de abril de 2018].

Matemáticas: Enseñanza Universitaria Vol. XIV Nº 2 Diciembre; (2006); Escalante Coteiro, C.; “*Distribuciones clase (a, b) y algoritmo de Panjer*” [Documento disponible en línea]. Recurso electrónico disponible en el siguiente enlace: <http://revistaerm.univalle.edu.co/VolXIV2/autor1.pdf> [Visitado el 3 de abril de 2018].

Ecos de Economía: A Latin American Journal of Applied Economics, Vol 14 No 30 (2010); Franco Arbeláez L.C.; Velasquez Ceballos, H.; “*Alternativas fundamentales para cuantificar el riesgo operacional*” [Documento disponible en línea]. Recurso electrónico disponible en el siguiente enlace: <http://publicaciones.eafit.edu.co/index.php/ecos-economia/article/view/178/194> [Visitado el 3 de abril de 2018].

Documentación y material académico

Balbás de la Corte, A.; “*Tarifificación de Seguros No Vida*”; (2017); Material didáctico facilitado para el seguimiento de la asignatura impartida por el profesor en el Máster en Ciencias Actuariales y Financieras de la Universidad Carlos III de Madrid.

Kalliopi, M.; “*Estadística Actuarial*”; (2016); Material didáctico facilitado para el seguimiento de la asignatura impartida por la profesora en el Máster en Ciencias Actuariales y Financieras de la Universidad Carlos III de Madrid.



Medina López, A.; “*Estadística Empresarial II*”; (2013); Material didáctico facilitado para el seguimiento de la asignatura impartida por la profesora en el Grado en Administración y Dirección de Empresas de la Universidad Rey Juan Carlos de Madrid.

Simón Luis, C.; “*Implementación Numérica de un Modelo en Seguros No Vida*”; (2012); Trabajo de Fin de Máster del Máster en Ciencias Actuariales y Financieras de la Universidad Carlos III de Madrid.

Recursos en internet

Universitat de Barcelona; “*La distribución Gamma*” [Documento disponible en línea]. Recurso electrónico disponible en el siguiente enlace: <http://www.ub.edu/stat/GrupsInnovacio/Statmedia/demo/Temas/Capitulo4/BOC4m1t7.htm> [Visitado el 10 de mayo de 2018].

Universitat de Valencia; Máster en Ciencias Actuariales y Financieras “*Modelos de Probabilidad II*” [Documento disponible en línea]. Recurso electrónico disponible en el siguiente enlace: <https://www.uv.es/lejarza/eaa/teoria/EAA6-b.pdf> [Visitado el 10 de mayo de 2018].

Semantic Scholar; Dorigoni; A. (2015); “*Inverse Gaussian Distribution*” [Documento disponible en línea]. Recurso electrónico disponible en el siguiente enlace: <https://pdfs.semanticscholar.org/e62d/5a35cf2527062f53884ed631dc2c810195a0.pdf> [Visitado el 13 de mayo de 2018].

William & Mary; Leemis, L. “*Log-Logistic distribution*” [Documento disponible en línea]. Recurso electrónico disponible en el siguiente enlace: <http://www.math.wm.edu/~leemis/chart/UDR/PDFs/Loglogistic.pdf> [Visitado el 17 de mayo de 2018].

Universitat de Barcelona; “*La distribución de Weibull*” [Documento disponible en línea]. Recurso electrónico disponible en el siguiente enlace: <http://www.ub.edu/stat/GrupsInnovacio/Statmedia/demo/Temas/Capitulo4/BOC4m1t9.htm> [Visitado el 18 de mayo de 2018].

Mathworks; “*Weibull cumulative distribution function*” [Recurso disponible en línea]. Recurso electrónico disponible en el siguiente enlace: <https://es.mathworks.com/help/stats/wblcdf.html> [Visitado el 18 de mayo de 2018].



Universitat de Barcelona; “*El p-valor*” [Documento disponible en línea]. Recurso electrónico disponible en el siguiente enlace:

<http://www.ub.edu/stat/GrupsInnovacio/Statmedia/demo/Temas/Capitulo9/BOC9m1t18.htm>

[Visitado el 20 de mayo de 2018].



ANEXOS



Anexo I. Código principal del Tarificador No Vida

```

%-----
%
% TRABAJO DE FIN DE MÁSTER
% MÉTODOS AVANZADOS PARA TARIFICACIÓN DE SEGUROS NO VIDA: APLICACIÓN DEL ALGORITMO
% RECURSIVO DE PANJER PARA LA CONSTRUCCIÓN DE UN MODELO ESTADÍSTICO DE PÉRDIDAS
% AGREGADAS
%
% Máster en Ciencias Actuariales y Financieras
% UNIVERSIDAD CARLOS III DE MADRID
%
&
% Alumno: Eugenio Margallo Pereira
%-----
function
ret=Tarificador_NoVida(ruta,m,alpha,c_ZM,h_span,simul,perc_costs,percentil,k1,k2)
    %Inputs de la función:

    %- ruta: será la ruta a la que se dirige el aplicativo para cargar
    %los datos almacenados en Excel.

    %- m: será el número total de clústeres en los que se ha dividido el
    %producto asegurador.

    %- alpha: Nivel de significación al que se realiza el estudio
    %estadístico para el ajuste de distribuciones de probabilidad.

    %- c_ZM: valor con el que se modifica la probabilidad de la función
    %en el punto 0 para las distribuciones de la clase (a,b,1).

    %- h_span: valor utilizado para la discretización de la
    %distribución de probabilidad continua para las severidades.

    %- simul: Número de veces que se realizarán las simulaciones para
    %evaluar la adecuación de las reservas.

    %- perc_costs: recargo sobre la severidad asociado a los costes
    %administrativos y de gestión de expedientes de siniestros.

    %percentil: percentil asociado al cálculo de las reservas y VaR.

    %- k1: recargo de seguridad sobre la esperanza del coste del
    %clúster.

    %- k2: recargo de riesgo sobre el VaR del clúster.

    %Introducción de BBDD en Excel del producto a analizar que contenga los
    %clústeres definidos previamente por la compañía
    %La estructura de la información estudiada se organiza de la siguiente
    %manera:

```




```

%Columna 1: Frecuencias de siniestros
%Columna 2: Severidades
%Columna 3: Identificación del clúster al que pertenece cada asegurado

data_mat=xlsread(ruta); res_cluster=data_mat;

%Generación de vector de nueve posiciones para almacenar la información que se
volcará
%finalmente en el aplicativo desarrollado (tablas de resultados y gráficos)
ret=cell(9);

%Tabla representación de ajustes de distribuciones para formulario
tabla1=cell(m,9);
%Tabla de representación de estadísticos y VaR para formulario
tabla2=cell(m,4);
%Tabla de tarificación del producto por cada clúster
tabla3=cell(m,7);
%Tablas para introducción de gráficos de frecuencias y severidades,
%en versión continua y discreta.
datos_plot_freqd=cell(m);datos_plot_freqc=cell(m);
datos_plot_sevc=cell(m);datos_plot_sevd=cell(m);
datos_grafico=cell(2);
datos_grafico{1}=[]; datos_grafico{2}=[];
%Tabla de resumen de la tarificación realizada
tabla4=cell(1,6);

for k=1:m %Bucle para el análisis individualizado de cada clúster

    tabla1{k,1}=k; tabla2{k,1}=k; tabla3{k,1}=k;

%
% AJUSTE DE LAS DISTRIBUCIONES DE FRECUENCIAS Y SEVERIDADES
%

%Identificación del clúster atendiendo a la Columna 3 de la Base de Datos
v=genvarname(['Cluster_',num2str(k)],who);
testfilter=res_cluster(:,3)==k;
submatrix1=res_cluster(testfilter,:);
eval([v ' = submatrix1']);

%Recuento del total de asegurados que pertenece al clúster analizado
long=genvarname(['long_cluster_',num2str(k)],who);
longit=length(submatrix1(:,1));
eval([long '=longit']);
[ncc(k)]=[longit];
tabla1{k,2}=longit; tabla3{k,2}=longit;

```



```

%-----
%      AJUSTE DE PARÁMETROS A DISTRIBUCIONES Y TEST CHI-CUADRADO
%      PARA LAS FRECUENCIAS SINIESTRALES
%-----

%Identificación del vector de las frecuencias de siniestros de cada asegurado
%que pertenece al clúster específico. Se identifica como la Columna 1 de la
%matriz de datos.
f = genvarname(['freq',num2str(k)],who);
    submatrix2=submatrix1(:,1);
eval([f '=submatrix2']);

resultado_tabulate=tabulate(submatrix2);
datos_plot_freqd{k}=cell(2);
datos_plot_freqd{k}{1}=resultado_tabulate(:,1)';
datos_plot_freqd{k}{2}=resultado_tabulate(:,3)'./100;

%DISTRIBUCIÓN BINOMIAL - CLASE (a,b,0)
%-----
%Proceso de ajuste de los parámetros por máxima verosimilitud de las
frecuencias del clúster a una distribución Binomial
BinomFit_freq = genvarname(['BinomFit_freq',num2str(k)],who);
    %Ajuste de los parámetros de la distribución
    submatfit_binom_freq=fitdist(submatrix2,'Binomial','NTrials',365); %Se
                                fija n=365 por los días del año
    param_n_bin_freq=submatfit_binom_freq.N; %Identificación del parámetro "n"
                                de la Binomial
    param_p_bin_freq=submatfit_binom_freq.p; %Identificación del parámetro "p"
                                de la Binomial

    %Proyección de probabilidades teóricas de la distribución Binomial según
    los parámetros obtenidos
    for i=1:(length(submatrix2)+1)
        theoprob_bin(i)=binopdf((i-1),param_n_bin_freq,param_p_bin_freq);
    end

    %Obtención del estadístico Chi-Cuadrado de la Binomial según su proyección
    teórica
    pdf_binom=length(submatrix2)*theoprob_bin;
eval([BinomFit_freq '=[submatfit_binom_freq]']);

% Comprobación mediante Test Chi-Cuadrado de que los datos siguen la
distribución Binomial con los parámetros estimados
BinomTest_freq=
    genvarname(['BinomTest_freq','Binom_pV_freq',num2str(k)],who);
    %Organización de datos para análisis de las frecuencias absolutas
    registradas en el clúster
    obs=tabulate(submatrix2);observ=obs(:,2);
    %Comprobación con Test Chi-Cuadrado de la distribución observada y teórica
    [submattest_binom_freq pV_binom_freq]=chisqr_test(observ,pdf_binom,alpha);

eval([BinomTest_freq '=[submattest_binom_freq pV_binom_freq]']);

```



```

%DISTRIBUCIÓN BINOMIAL "ZM" y "ZT" - CLASE (a,b,1)
%-----
%Cálculo de probabilidades para la distribución modificada en cero ("ZM")
Binom_freq_ZM = genvarname(['Binom_ZM_test_', 'Binom_ZM_pV_', num2str(k)], who);
    param_n_binom_freq=submatfit_binom_freq.N; %Identificación del parámetro
        "n" de la Binomial
    param_p_binom_freq=submatfit_binom_freq.p; %Identificación del parámetro
        "p" de la Binomial

%Se procede a la proyección de probabilidades aplicando una modificación
en 0 que toma el valor c_ZM.
theo_bino=zero_mod_bino(param_n_binom_freq,param_p_binom_freq,c_ZM);
%Obtención del estadístico Chi-Cuadrado de la Binomial ZM según su
proyección teórica
theo_bino=length(submatrix2)*theo_bino(1, [1:length(observ)]);
%Comprobación con Test Chi-Cuadrado de la distribución observada y teórica
[test_binom_ZM pV_binom_ZM]=chisqr_test(observ,theo_bino,alpha);
eval([Binom_freq_ZM '=[test_binom_ZM pV_binom_ZM]']);

%Cálculo de probabilidades para la distribución truncada en cero ("ZT")
Binom_freq_ZT=genvarname(['Binom_ZT_test_', 'Binom_ZT_pV_', num2str(k)], who);
    %Se procede a la proyección de probabilidades truncando la distribución en
    0
    theo_binom=zero_trun_bino(param_n_binom_freq,param_p_binom_freq);
    %Obtención del estadístico Chi-Cuadrado de la Binomial ZT según su
    proyección teórica
    theo_binom=length(submatrix2)*theo_binom(1, [1:length(observ)]);
    %Comprobación con Test Chi-Cuadrado de la distribución observada y teórica
    [test_binom_ZT pV_binom_ZT]=chisqr_test(observ,theo_binom,alpha);
eval([Binom_freq_ZT '=[test_binom_ZT pV_binom_ZT]']);

%DISTRIBUCIÓN BINOMIAL NEGATIVA - CLASE (a,b,0)
%-----
%Proceso de ajuste de los parámetros por máxima verosimilitud de las
frecuencias del clúster a una distribución Binomial Negativa
NegBinomFit_freq = genvarname(['NegBinomFit_freq_', num2str(k)], who);
test_negbinom=nbinfit(submatrix2); %Ajuste de los parámetros de la
distribución para filtrado

if test_negbinom==[Inf 1]
    submatfit_negbinom_freq=0; %Filtro para evitar errores en caso de no
        poder realizar el ajuste correcto
    param_R_negbin_freq=0; %Identificación del parámetro "r" de la
        Binomial Negativa
    param_P_negbin_freq=0; %Identificación del parámetro "p" de la
        Binomial Negativa

Else

%Ajuste de los parámetros de la distribución
submatfit_negbinom_freq = fitdist(submatrix2, 'NegativeBinomial');
param_R_negbin_freq=submatfit_negbinom_freq.R; %Identificación del
        arámetro "r" de la Binomial Negativa
    
```



```

param_P_negbin_freq=submatfit_negbinom_freq.P; %Identificación del
                                parámetro "p" de la Binomial Negativa
%Proyección de probabilidades teóricas de la distribución Binomial
Negativa según los parámetros obtenidos
    for i=1:(length(submatrix2)+1)
        theoprob_negbinom(i)=nbinpdf((i-
                                1),param_R_negbin_freq,param_P_negbin_freq);
    end
%Obtención del estadístico Chi-Cuadrado de la Binomial Negativa según
su proyección teórica
pdf_negbinom=length(submatrix2)*theoprob_negbinom;
end
eval([NegBinomFit_freq 'submatfit_negbinom_freq']);

% Comprobación mediante Test Chi-Cuadrado de que los datos siguen la
distribución Binomial Negativa con los parámetros estimados
NegBinomTest_freq =
    genvarname(['NegBinomTest_freq_', 'NegBinom_pV_freq_', num2str(k)], who);
if test_negbinom~=[Inf 1]
    [submattest_negbinom_freq
    pV_negbinom_freq]=chisqr_test(observ, pdf_negbinom, alpha);
else if test_negbinom==[Inf 1]
    [submattest_negbinom_freq]=1; [pV_negbinom_freq]=0; %Caso de no poder
                                realizar el cálculo
end
end
eval([NegBinomTest_freq 'submattest_negbinom_freq pV_negbinom_freq']);

%DISTRIBUCIÓN BINOMIAL NEGATIVA "ZM" y "ZT" - CLASE (a,b,1)
%-----
%Cálculo de probabilidades para la distribución modificada en cero ("ZM")

NegBino_freq_ZM =
    genvarname(['NegBin_ZM_test_', 'NegBin_ZM_pV_', num2str(k)], who);
if test_negbinom==[Inf 1]
    [test_nbino_ZM]=1, [pV_nbino_ZM]=0;
eval([NegBino_freq_ZM 'test_nbino_ZM pV_nbino_ZM']);
else if test_negbinom~=[Inf 1]
    param_R_nbino_freq=submatfit_negbinom_freq.R; %Identificación del
                                parámetro "r" de la Binomial Negativa
    param_P_nbino_freq=submatfit_negbinom_freq.P; %Identificación del
                                parámetro "p" de la Binomial Negativa
%Organización de datos para análisis de las frecuencias absolutas
registradas en el clúster
obs=tabulate(submatrix2); observ=obs(:,2); n=max(obs(:,1));
%Se procede a la proyección de probabilidades aplicando una
modificación en 0 que toma el valor c_ZM.
theo_nbino=zero_mod_nbin(param_R_nbino_freq,param_P_nbino_freq,c_ZM,n);
%Obtención del estadístico Chi-Cuadrado de la Binomial Negativa ZM
según su proyección teórica

```



```

        theo_nbino=length(submatrix2)*theo_nbino(1,[1:length(observ)]);
        %Comprobación con Test Chi-Cuadrado de la distribución observada y
        teórica
        [test_nbino_ZM, pV_nbino_ZM]=chisqr_test(observ,theo_nbino,alpha);
        eval(['NegBino_freq_ZM ']=[test_nbino_ZM pV_nbino_ZM]);
        end
    end

    %Cálculo de probabilidades para la distribución truncada en cero
    NegBin_freq_ZT =
        genvarname(['NegBin_ZT_test_', 'NegBin_ZT_pV_', num2str(k)], who);
    if test_negbinom==[Inf 1]
        [test_nbino_ZT]=1, [pV_nbino_ZT]=0;
        eval(['NegBin_freq_ZT ']=[test_nbino_ZT pV_nbino_ZT]);
    else
        param_R_nbinom_freq=submatfit_negbinom_freq.R;%Identificación del
        parámetro "r" de la Binomial Negativa
        param_P_nbinom_freq=submatfit_negbinom_freq.P;%Identificación del
        parámetro "p" de la Binomial Negativa
        %Organización de datos para análisis de las frecuencias absolutas
        registradas en el clúster
        obs=tabulate(submatrix2); observ=obs(:,2); n=max(obs(:,1));
        %Se procede a la proyección de probabilidades truncando la distribución en
        0
        theo_nbinom=zero_trun_nbin(param_R_nbinom_freq,param_P_nbinom_freq,n);
        %Obtención del estadístico Chi-Cuadrado de la Binomial Negativa ZT según
        su proyección teórica
        theo_nbinom=length(submatrix2)*theo_nbinom(1,[1:length(observ)]);
        %Comprobación con Test Chi-Cuadrado de la distribución observada y teórica
        [test_nbino_ZT pV_nbino_ZT]=chisqr_test(observ,theo_nbinom,alpha);
        eval(['NegBin_freq_ZT ']=[test_nbino_ZT pV_nbino_ZT]);
    end

    %DISTRIBUCIÓN GEOMÉTRICA - CLASE (a,b,0)
    %-----
    %Proceso de ajuste de los parámetros por máxima verosimilitud de las
    frecuencias del
    %clúster a una distribución Geométrica
    GeometFit_freq=genvarname(['GeometFit_freq_', num2str(k)], who);
    %Identificación del parámetro "p" de la Geométrica
    submatfit_Geomet_freq=mle(submatrix2, 'distribution', 'Geometric');
    %Proyección de probabilidades teóricas de la distribución Geométrica según
    los parámetros obtenidos

        for i=1:(length(submatrix2)+1)
            theoprob_geo(i)=geopdf((i-1), submatfit_Geomet_freq);
        end

    %Obtención del estadístico Chi-Cuadrado de la Geométrica según su
    proyección teórico
    pdf_geo=length(submatrix2)*theoprob_geo(1,[1:length(observ)]);
    eval(['GeometFit_freq ']=submatfit_Geomet_freq);

```



```

% Comprobación mediante Test Chi-Cuadrado de que los datos siguen la
distribución Geométrica con
% los parámetros estimados
GeometTest_freq =
    genvarname(['GeometTest_freq_', 'Geomet_pV_freq_', num2str(k)], who);
    %Organización de datos para análisis de las frecuencias absolutas
    registradas en el clúster
    obs=tabulate(submatrix2); observ=obs(:,2); n=max(submatrix2);
    %Comprobación con Test Chi-Cuadrado de la distribución observada y teórica
    [submattest_Geomet_freq pV_Geomet_freq]=chisqr_test(observ,pdf_geo,alpha);
eval(['GeometTest_freq ']=[submattest_Geomet_freq pV_Geomet_freq]);

%DISTRIBUCIÓN GEOMÉTRICA "ZM" y "ZT" - CLASE (a,b,1)
%-----
%Cálculo de probabilidades para la distribución modificada en cero ("ZM")

Geomet_freq_ZM=genvarname(['Geomet_ZM_test_', 'Geomet_ZM_pV_', num2str(k)], who);
    param_p_geomet_freq=submatfit_Geomet_freq; %Identificación del parámetro
    "p" de la Geométrica
    %Organización de datos para análisis de las frecuencias absolutas
    registradas en el clúster
    obs=tabulate(submatrix2); observ=obs(:,2); n=max(observ(:,1));
    %Se procede a la proyección de probabilidades modificando la distribución
    en 0
    theo_geom=zero_mod_geom(param_p_geomet_freq,c_ZM,n);
    %Obtención del estadístico Chi-Cuadrado de la Geométrica ZM según su
    proyección teórica
    theo_geom=length(submatrix2)*theo_geom(1, [1:length(observ)]);
    %Comprobación con Test Chi-Cuadrado de la distribución observada y teórica
    [test_geomet_ZM pV_geomet_ZM]=chisqr_test(observ,theo_geom,alpha);
eval(['Geomet_freq_ZM ']=[test_geomet_ZM pV_geomet_ZM]);

%Cálculo de probabilidades para la distribución modificada en cero ("ZT")
Geomet_freq_ZT=genvarname(['Geomet_ZT_test_', 'Geomet_ZT_pV_', num2str(k)], who);
    %Se procede a la proyección de probabilidades truncando la distribución en
    0
    theo_geome=zero_trun_geom(param_p_geomet_freq,n);
    %Obtención del estadístico Chi-Cuadrado de la Geométrica ZT según su
    proyección teórica
    theo_geome=length(submatrix2)*theo_geome(1, [1:length(observ)]);
    %Comprobación con Test Chi-Cuadrado de la distribución observada y teórica
    [test_geomet_ZT pV_geomet_ZT]=chisqr_test(observ,theo_geome,alpha);

eval(['Geomet_freq_ZT ']=[test_geomet_ZT pV_geomet_ZT]);

%DISTRIBUCIÓN DE POISSON - CLASE (a,b,0)
%-----
%Proceso de ajuste de los parámetros por máxima verosimilitud de las
frecuencias del clúster a una distribución Poisson

```



```

PoissonFit_freq = genvarname(['PoissonFit_freq_', num2str(k)], who);
    %Ajuste de los parámetros de la distribución
    submatfit_Poisson_freq=fitdist(submatrix2, 'Poisson');
    param_lambda_poiss_freq=submatfit_Poisson_freq.lambda; %Identificación del
    parámetro "lambda" de Poisson
    %Proyección de probabilidades teóricas de la distribución de Poisson según
    los parámetros obtenidos
    for i=1:(length(submatrix2)+1)
        theoprob_poiss(i)=poisspdf((i-1), param_lambda_poiss_freq);
    end
    %Obtención del estadístico Chi-Cuadrado de la Geométrica según su
    proyección teórica
    pdf_poiss=length(submatrix2)*theoprob_poiss;
eval([PoissonFit_freq '=submatfit_Poisson_freq']);

% Comprobación mediante el test Chi-Cuadrado de que los parámetros
% obtenidos siguen la distribución específica (alfa=0.05)
PoissonTest_freq =
    genvarname(['PoissonTest_freq_', 'Poisson_pV_freq_', num2str(k)], who);
    %Organización de datos para análisis de las frecuencias absolutas
    registradas en el clúster
    obs=tabulate(submatrix2); observ=obs(:,2); n=max(submatrix2);
    %Comprobación con Test Chi-Cuadrado de la distribución observada y teórica
    [submattest_Poisson_freq
    pV_Poisson_freq]=chisqr_test(observ, pdf_poiss, alpha);
eval([PoissonTest_freq '=[submattest_Poisson_freq pV_Poisson_freq]']);

%DISTRIBUCIÓN DE POISSON "ZM" y "ZT" - CLASE (a,b,1)
%-----
%Cálculo de probabilidades para la distribución modificada en cero ("ZM")

%Cálculo de probabilidades para la distribución modificada en cero
Poisson_freq_ZM =
    genvarname(['Poisson_ZM_test_', 'Poisson_ZM_pV_', num2str(k)], who);
    param_lambda_poisson_freq=submatfit_Poisson_freq.lambda;
    %Organización de datos para análisis de las frecuencias absolutas
    registradas en el clúster
    obs=tabulate(submatrix2); observ=obs(:,2); n=max(obs(:,1));
    %Se procede a la proyección de probabilidades modificando la distribución
    en 0

    theo_pois=zero_mod_poiss(param_lambda_poisson_freq, c_ZM, n);
    %Obtención del estadístico Chi-Cuadrado de Poisson ZT según su proyección
    teórica
    theo_pois=length(submatrix2)*theo_pois(1, [1:length(observ)]);
    %Comprobación con Test Chi-Cuadrado de la distribución observada y teórica
    [test_poisson_ZM pV_poisson_ZM]=chisqr_test(observ, theo_pois, alpha);
eval([Poisson_freq_ZM '=[test_poisson_ZM pV_poisson_ZM]']);

```



```

%Cálculo de probabilidades para la distribución truncada en cero
Poisson_freq_ZT=genvarname(['Poisson_ZT_test_', 'Poisson_ZT_pV_', num2str(k)], who);
    %Se procede a la proyección de probabilidades truncando la distribución en
    0
    theo_poiss=zero_trun_poiss(param_lambda_poisson_freq,n);
    %Obtención del estadístico Chi-Cuadrado de Poisson ZT según su proyección
    teórica
    theo_poiss=length(submatrix2)*theo_poiss(1, [1:length(observ)]);
    %Comprobación con Test Chi-Cuadrado de la distribución observada y teórica
    [test_poisson_ZT pV_poisson_ZT]=chisqr_test(observ,theo_poiss,alpha);
eval(['Poisson_freq_ZT ']=[test_poisson_ZT pV_poisson_ZT]);

%-----
%      AJUSTE DE PARÁMETROS A DISTRIBUCIONES Y TEST CHI-CUADRADO
%      PARA LAS SEVERIDADES
%-----
%Extracción de las severidades de cada asegurado que pertenece al
%clúster específico
s = genvarname(['sev', num2str(k)], who);
    submatrix3=submatrix1(:,2);
eval([s '=submatrix3']);

%DISTRIBUCIÓN EXPONENCIAL
%-----
%Proceso de ajuste de los parámetros por máxima verosimilitud de las
frecuencias del
%clúster a una distribución Exponencial
ExpFit_sev = genvarname(['ExpFit_sev_', num2str(k)], who);
    %Ajuste de los parámetros de la distribución
    submatfit_exp_sev=fitdist(submatrix3, 'Exponential');
    param_mu_exp_sev=submatfit_exp_sev.mu; %Identificación del parámetro "mu"
    de la Exponencial
eval(['ExpFit_sev '=submatfit_exp_sev']);

%Comprobación con Test Chi-Cuadrado de la distribución observada y teórica
ExpTest_sev = genvarname(['ExpTest_sev_', 'Exp_pV_sev_', num2str(k)], who);
%Aplicación del Test Chi-Cuadrado para determinar el resultado ajuste
    [submattest_exp_sev

        pV_exp_sev]=chi2gof(submatrix3, 'CDF', submatfit_exp_sev, 'Alpha', alpha);
eval(['ExpTest_sev ']=[submattest_exp_sev pV_exp_sev]);

%DISTRIBUCIÓN GAMMA
%-----
%Proceso de ajuste de los parámetros por máxima verosimilitud de las
frecuencias del clúster a una distribución Gamma
GammaFit_sev = genvarname(['GammaFit_sev_', num2str(k)], who);
    %Ajuste de los parámetros de la distribución
    submatfit_gamma_sev=fitdist(submatrix3, 'Gamma');
    param_a_gamma_sev=submatfit_gamma_sev.a; %Identificación del parámetro "a"
    de la Gamma

```




```

param_b_gamma_sev=submatfit_gamma_sev.b;%Identificación del parámetro "b"
                                de la Gamma
eval([GammaFit_sev '=submatfit_gamma_sev']);

%Comprobación con Test Chi-Cuadrado de la distribución observada y teórica
GammaTest_sev = genvarname(['GammaTest_sev_', 'Gamma_pV_sev_', num2str(k)], who);
%Aplicación del Test Chi-Cuadrado para determinar el resultado ajuste
[submattest_gamma_sev pV_gamma_sev]
    =chi2gof(submatrix3, 'CDF', submatfit_gamma_sev, 'Alpha', alpha);
eval([GammaTest_sev '=[submattest_gamma_sev pV_gamma_sev]']);

%DISTRIBUCIÓN GENERALIZED-PARETO
%-----
%Proceso de ajuste de los parámetros por máxima verosimilitud de las
frecuencias del clúster a una distribución GeneralizedPareto
Pareto_filter='GPareto';
GParetoFit_sev=genvarname(['GParetoFit_sev_', num2str(k)], who);
if Pareto_filter=='GPareto';
    %Eliminación de valores 0 para poder efectuar el ajuste
    pos=find(submatrix3(:,1)<1); submatrix3(pos,:)=1;
    %Ajuste de los parámetros de la distribución
    submatfit_gpareto_sev=fitdist(submatrix3, 'GeneralizedPareto');
    param_k_gpareto_sev=submatfit_gpareto_sev.k; %Identificación del
                                parámetro "k" de la GPareto
    param_sigma_gpareto_sev=submatfit_gpareto_sev.sigma;%Identificación
                                del parámetro "sigma" de la GPareto
    param_theta_gpareto_sev=submatfit_gpareto_sev.theta;%Identificación
                                del parámetro "theta" de la GPareto
    if submatfit_gpareto_sev.k<-0.5; %Casuística de caso en el que no
                                puede efectuarse el cálculo
        'Cálculo no adecuado';
    end
end
eval([GParetoFit_sev '=submatfit_gpareto_sev']);

%Comprobación con Test Chi-Cuadrado de la distribución observada y teórica
GParetoTest_sev =
    genvarname(['GParetoTest_sev_', 'GPareto_pV_sev_', num2str(k)], who);

if submatfit_gpareto_sev.k>=-0.5
    %Aplicación del Test Chi-Cuadrado para determinar el resultado ajuste
    [submattest_gpareto_sev
pV_gpareto_sev]=chi2gof(submatrix3, 'CDF', submatfit_gpareto_sev, 'Alpha', alpha);
else 'Cálculo no adecuado';
    [submattest_gpareto_sev]=1 ; [pV_gpareto_sev]=0;
end
eval([GParetoTest_sev '=[submattest_gpareto_sev pV_gpareto_sev]']);

```



```

%DISTRIBUCIÓN INVERSA GAUSSIANA
%-----
%Proceso de ajuste de los parámetros por máxima verosimilitud de las
frecuencias del clúster a una distribución Gamma
InvGaussFit_sev = genvarname(['InvGaussFit_sev_',num2str(k)],who);
    %Ajuste de los parámetros de la distribución
    submatfit_InvGauss_sev=fitdist(submatrix3,'InverseGaussian');
    param_mu_InvGauss_sev=submatfit_InvGauss_sev.mu; %Identificación del
                                parámetro "mu" de la Inversa Gaussiana
    param_lambda_InvGauss_sev=submatfit_InvGauss_sev.lambda; %Identificación
                                del parámetro "lambda" de la Inversa Gaussiana
eval([InvGaussFit_sev '=submatfit_InvGauss_sev']);

%Comprobación con Test Chi-Cuadrado de la distribución observada y teórica
InvGaussTest_sev =
    genvarname(['InvGaussTest_sev_','InvGauss_pV_sev_',num2str(k)],who);
    %Aplicación del Test Chi-Cuadrado para determinar el resultado ajuste
    [submattest_InvGauss_sev pV_InvGauss_sev]
        =chi2gof(submatrix3,'CDF',submatfit_InvGauss_sev,'Alpha',alpha);
eval([InvGaussTest_sev '=[submattest_InvGauss_sev pV_InvGauss_sev]']);

%DISTRIBUCIÓN LOGISTIC
%-----
%Proceso de ajuste de los parámetros por máxima verosimilitud de las
frecuencias del clúster a una distribución Logistic
LogiFit_sev = genvarname(['LogiFit_sev_',num2str(k)],who);
    %Ajuste de los parámetros de la distribución
    submatfit_Logi_sev=fitdist(submatrix3,'Logistic');
    param_mu_Logi_sev=submatfit_Logi_sev.mu; %Identificación del parámetro
                                "mu" de la Logistic
    param_sigma_Logi_sev=submatfit_Logi_sev.sigma; %Identificación del
                                parámetro "sigma" de la Logistic
eval([LogiFit_sev '=submatfit_Logi_sev']);
%Comprobación con Test Chi-Cuadrado de la distribución observada y teórica
LogiTest_sev = genvarname(['LogiTest_sev_','Logi_pV_sev_',num2str(k)],who);
%Aplicación del Test Chi-Cuadrado para determinar el resultado ajuste
[submattest_Logi_sev
    pV_Logi_sev]=chi2gof(submatrix3,'CDF',submatfit_Logi_sev,'Alpha',alpha);
eval([LogiTest_sev '=[submattest_Logi_sev pV_Logi_sev]']);

%DISTRIBUCIÓN LOG-LOGISTIC
%-----
%Proceso de ajuste de los parámetros por máxima verosimilitud de las
frecuencias del clúster a una distribución LogLogistic
LogLogi_filter='LogLogi';
LogLogiFit_sev=genvarname(['LogLogi_sev_',num2str(k)],who);
if LogLogi_filter=='LogLogi';
    %Eliminación de valores 0 para poder efectuar el ajuste
    pos=find(submatrix3(:,1)<1); submatrix3(pos,:)=1;
    %Ajuste de los parámetros de la distribución
    submatfit_LogLogi_sev=fitdist(submatrix3,'LogLogistic');

```



```

param_mu_LogLogi_sev=submatfit_LogLogi_sev.mu; %Identificación del
                                parámetro "mu" de la LogLogistic
param_sigma_LogLogi_sev=submatfit_LogLogi_sev.sigma; %Identificación del
                                parámetro "sigma" de la LogLogistic
eval([LogLogiFit_sev '=[submatfit_LogLogi_sev]']);
end
%Comprobación con Test Chi-Cuadrado de la distribución observada y teórica
LogLogiTest_sev =
    genvarname(['LogLogiTest_sev_', 'LogLogi_pV_sev_', num2str(k)], who);
    %Aplicación del Test Chi-Cuadrado para determinar el resultado ajuste
    [submattest_LogLogi_sev
pV_LogLogi_sev]=chi2gof(submatrix3, 'CDF', submatfit_LogLogi_sev, 'Alpha', alpha);
eval([LogLogiTest_sev '=[submattest_LogLogi_sev pV_LogLogi_sev]']);

%DISTRIBUCIÓN LOG-NORMAL
%-----
%Proceso de ajuste de los parámetros por máxima verosimilitud de las
frecuencias del clúster a una distribución LogNormal
LogNorm_filter='LogNorm';
LogNormFit_sev=genvarname(['LogNorm_sev_', num2str(k)], who);
    if LogNorm_filter=='LogNorm'
        %Eliminación de valores 0 para poder efectuar el ajuste
        pos=find(submatrix3(:,1)<1); submatrix3(pos,:)=1;
        %Ajuste de los parámetros de la distribución
        submatfit_LogNorm_sev=fitdist(submatrix3, 'LogNormal');
        param_mu_LogNorm_sev=submatfit_LogNorm_sev.mu; %Identificación del
                                parámetro "mu" de la LogNormal
        param_sigma_LogNorm_sev=submatfit_LogNorm_sev.sigma; %Identificación
                                del parámetro "sigma" de la LogNormal
        eval([LogNormFit_sev '=[submatfit_LogNorm_sev]']);
    end
%Comprobación con Test Chi-Cuadrado de la distribución observada y teórica
LogNormTest_sev =
    genvarname(['LogNormTest_sev_', 'LogNorm_pV_sev_', num2str(k)], who);
    %Aplicación del Test Chi-Cuadrado para determinar el resultado ajuste
    [submattest_LogNorm_sev
pV_LogNorm_sev]=chi2gof(submatrix3, 'CDF', submatfit_LogNorm_sev, 'Alpha', alpha);

eval([LogNormTest_sev '=[submattest_LogNorm_sev pV_LogNorm_sev]']);

%DISTRIBUCIÓN NORMAL
%-----
%Proceso de ajuste de los parámetros por máxima verosimilitud de las
frecuencias del clúster a una distribución Normal
NormFit_sev = genvarname(['NormFit_sev_', num2str(k)], who);
    %Ajuste de los parámetros de la distribución
    submatfit_Norm_sev=fitdist(submatrix3, 'Normal');
    param_mu_Norm_sev=submatfit_Norm_sev.mu; %Identificación del parámetro
                                "mu" de la Normal
    param_sigma_Norm_sev=submatfit_Norm_sev.sigma; %Identificación del
                                parámetro "sigma" de la Normal

```



```

eval([NormFit_sev '=submatfit_Norm_sev']);
%Comprobación con Test Chi-Cuadrado de la distribución observada y teórica
NormTest_sev = genvarname(['NormTest_sev_', 'Norm_pV_sev_', num2str(k)], who);
    %Aplicación del Test Chi-Cuadrado para determinar el resultado ajuste
    [submatteest_Norm_sev
        pV_Norm_sev]=chi2gof(submatrix3, 'CDF', submatfit_Norm_sev, 'Alpha', alpha);
eval([NormTest_sev '=[submatteest_Norm_sev pV_Norm_sev]']);

%DISTRIBUCIÓN WEIBULL
%-----
%Proceso de ajuste de los parámetros por máxima verosimilitud de las
severidades del clúster a una distribución Weibull
Weibull_filter='Weibull';
WeibullFit_sev=genvarname(['Weibull_sev_', num2str(k)], who);
    if Weibull_filter=='Weibull';
        %Eliminación de valores 0 para poder efectuar el ajuste
        pos=find(submatrix3(:,1)<1); submatrix3(pos,:)=1;
        %Ajuste de los parámetros de la distribución
        submatfit_Weibull_sev=fitdist(submatrix3, 'Weibull');
        param_A_Weibull_sev=submatfit_Weibull_sev.A %Identificación del
            parámetro "A" de la Weibull
        param_B_Weibull_sev=submatfit_Weibull_sev.B %Identificación del
            parámetro "B" de la Weibull
        eval([WeibullFit_sev '=[submatfit_Weibull_sev]']);
    end
    %Comprobación con Test Chi-Cuadrado de la distribución observada y teórica
    WeibullTest_sev =
        genvarname(['WeibullTest_sev_', 'Weibull_pV_sev_', num2str(k)], who);
        %Aplicación del Test Chi-Cuadrado para determinar el resultado ajuste
        [submatteest_Weibull_sev
            pV_Weibull_sev]=chi2gof(submatrix3, 'CDF', submatfit_Weibull_sev, 'Alpha', alpha);
        eval([WeibullTest_sev '=[submatteest_Weibull_sev pV_Weibull_sev]']);

%
%-----
%
%           SELECCIÓN DE LA DISTRIBUCIÓN DE PROBABILIDAD QUE MEJOR SE AJUSTA AL
%           CONJUNTO DE DATOS DE CADA CLUSTER ANALIZADO
%-----
%

%Codificación de las distribuciones para poder extraer la que mejor
%características presenta para la distribución estudiada:

    %1.1 Binomial, 1.2 Binomial ZM,, 1.3 Binomial ZT,
    %2.1 Binomial Neg, 2.2 Binomial Neg ZM 2.3 Binomial Neg ZT
    %3.1 Geométrica, 3.2 Geométrica ZM, 3.3 Geométrica ZT
    %4.1 Poisson, 4.2 Poisson ZM, 4.3 Poisson ZT

    Mat_Fit_Test_freq=genvarname(['Mat_Fit_Test_freq_cluster_', num2str(k)], who);
    mat_resumen_freq=[longit 0 0 0 0;
        1.1 submatteest_binom_freq, round(pV_binom_freq,2)
        round(param_n_binom_freq,2) round(param_p_binom_freq,2);

```

```

        1.2 test_binom_ZM, round(pV_binom_ZM,2)
round(param_n_binom_freq,2) round(param_p_binom_freq,2);
        1.3 test_binom_ZT, round(pV_binom_ZT,2)
round(param_n_binom_freq,2) round(param_p_binom_freq,2);
        2.1 submattest_negbinom_freq, round(pV_negbinom_freq,2)
round(param_R_negbin_freq,2) round(param_P_negbin_freq,2);
        2.2 test_nbino_ZM, round(pV_nbino_ZM,2)
round(param_R_negbin_freq,2) round(param_P_negbin_freq,2);
        2.3 test_nbino_ZT, round(pV_nbino_ZT,2)
round(param_R_negbin_freq,2) round(param_P_negbin_freq,2);
        3.1 submattest_Geomet_freq, round(pV_Geomet_freq,2)
round(submatfit_Geomet_freq,2) 0;
        3.2 test_geomet_ZM, round(pV_geomet_ZM,2)
round(submatfit_Geomet_freq,2) 0;
        3.3 test_geomet_ZT, round(pV_geomet_ZT,2)
round(submatfit_Geomet_freq,2) 0;
        4.1 submattest_Poisson_freq, round(pV_Poisson_freq,2)
round(param_lambda_poiss_freq,2) 0;
        4.2 test_poisson_ZM, round(pV_poisson_ZM,2)
round(param_lambda_poiss_freq,2) 0;
        4.3 test_poisson_ZT, round(pV_poisson_ZT,2)
round(param_lambda_poiss_freq,2) 0];
    format long g
    eval([Mat_Fit_Test_freq '=mat_resumen_freq'])
    %Comandos para seleccionar la mejor opción de cada clúster según la tabla
    resumen anterior
    Freq_dist_cluster=genvarname(['Freq_dist_cluster_',num2str(k)],who);
    max_pv_freq=max(mat_resumen_freq(2:13,3)); [row1
    col1]=find(mat_resumen_freq==max_pv_freq);
    [freq_dist]=mat_resumen_freq(row1,[1 4 5]);
    eval([Freq_dist_cluster '=[freq_dist]']);

    %Asignación del nombre de la distribución de frecuencias respecto de la
    %codificación anterior para identificar en tabla del formulario

    datos_plot_freqc{k}=cell(2);
    datos_plot_freqc{k}{1}=[0];
    datos_plot_freqc{k}{2}=[0];

    if freq_dist(1)==1.1
        %Gráfico de Binomial de la clase (a,b,0)
        tabla1{k,3}='Binomial (n, p)';
        freqmaxima=max(datos_plot_freqd{k}{1});
        freq_x=0:freqmaxima;
        for h=1:(freqmaxima+1)
            [vect(h)]=binopdf((h-1),freq_dist(2),freq_dist(3));
        end
        freq_y=[vect]; freq_y=freq_y(1:freqmaxima+1);
        datos_plot_freqc{k}{2}=freq_y;
        datos_plot_freqc{k}{1}=freq_x;
    end

```

```

elseif freq_dist(1)==1.2
    %Gráfico de la Binomial ZM de la clase (a,b,1)
    tablal{k,3}='Binomial ZM (n, p)';
    freqmaxima=max(datos_plot_freqd{k}{1});
    freq_x=0:freqmaxima;
    freq_y=zero_mod_bino(freq_dist(2),freq_dist(3),c_ZM);
    freq_y=freq_y(1:freqmaxima+1);
    datos_plot_freqc{k}{2}=freq_y;
    datos_plot_freqc{k}{1}=freq_x;
elseif freq_dist(1)==1.3
    %Gráfico de la Binomial ZT de la clase (a,b,1)
    tablal{k,3}='Binomial ZT (n, p)';
    freqmaxima = max(datos_plot_freqd{k}{1});
    freq_x=0:freqmaxima;
    freq_y=zero_trun_bino(freq_dist(2),freq_dist(3));
    freq_y=freq_y(1:freqmaxima+1);
    datos_plot_freqc{k}{2} = freq_y;
    datos_plot_freqc{k}{1} = freq_x;
elseif freq_dist(1)==2.1
    %Gráfico de la Binomial Negativa de la clase (a,b,0)
    tablal{k,3}='Bin. Negativa (r, p)';
    freqmaxima=max(datos_plot_freqd{k}{1});
    freq_x=0:freqmaxima;
    for h=1:(freqmaxima+1)
        [vect(h)]=nbinpdf((h-1),freq_dist(2),freq_dist(3));
    end
    freq_y=[vect]; freq_y=freq_y(1:freqmaxima+1);
    datos_plot_freqc{k}{2}=freq_y;
    datos_plot_freqc{k}{1}=freq_x;
elseif freq_dist(1)==2.2
    %Gráfico de la Binomial Negativa ZM de la clase (a,b,1)
    tablal{k,3}='Bin. Negativa ZM (r, p)';
    freqmaxima=max(datos_plot_freqd{k}{1});
    freq_x=0:freqmaxima;
    freq_y=zero_mod_nbin(freq_dist(2),freq_dist(3),c_ZM,freqmaxima);
    freq_y=[freq_y(1:freqmaxima+1)];
    datos_plot_freqc{k}{2}=freq_y;
    datos_plot_freqc{k}{1}=freq_x;
elseif freq_dist(1)==2.3
    %Gráfico de la Binomial Negativa ZT de la clase (a,b,1)
    tablal{k,3}='Bin. Negativa ZT (r, p)';
    freqmaxima=max(datos_plot_freqd{k}{1});
    freq_x=0:freqmaxima;
    freq_y=zero_trun_nbin(freq_dist(2),freq_dist(3),freqmaxima);
    freq_y=freq_y(1:freqmaxima+1);
    datos_plot_freqc{k}{2}=freq_y;
    datos_plot_freqc{k}{1}=freq_x;
elseif freq_dist(1)==3.1
    %Gráfico de la Geométrica de la clase (a,b,0)
    tablal{k,3}='Geométrica (p)';

```

```

    freqmaxima=max(datos_plot_freqd{k}{1});
    freq_x=0:freqmaxima;
    for h=1:(freqmaxima+1)
        [vect(h)]=geopdf((h-1),freq_dist(2));
    end
    freq_y=[vect]; freq_y=freq_y(1:freqmaxima+1);
    datos_plot_freqc{k}{2}=freq_y;
    datos_plot_freqc{k}{1}=freq_x;
elseif freq_dist(1)==3.2
    %Gráfico de la Geométrica ZM de la clase (a,b,1)
    tabla1{k,3}='Geométrica ZM (p)';
    freqmaxima=max(datos_plot_freqd{k}{1});
    freq_x=0:freqmaxima;
    freq_y=zero_mod_geo(freq_dist(2),c_ZM,freqmaxima);
    freq_y=freq_y(1:freqmaxima+1);
    datos_plot_freqc{k}{2}=freq_y;
    datos_plot_freqc{k}{1}=freq_x;
elseif freq_dist(1)==3.3
    %Gráfico de la Geométrica ZT de la clase (a,b,1)
    tabla1{k,3}='Geométrica ZT (p)';
    freqmaxima=max(datos_plot_freqd{k}{1});
    freq_x=0:freqmaxima;
    freq_y=zero_trun_geo(freq_dist(2),freqmaxima);
    freq_y=freq_y(1:freqmaxima+1);
    datos_plot_freqc{k}{2}=freq_y;
    datos_plot_freqc{k}{1}=freq_x;
elseif freq_dist(1)==4.1
    %Gráfico de Poisson de la clase (a,b,0)
    tabla1{k,3}='Poisson (lambda)';

    freqmaxima=max(datos_plot_freqd{k}{1});
    freq_x=0:freqmaxima;
    for h=1:(freqmaxima+1)
        [vect(h)]=poisspdf((h-1),freq_dist(2));
    end
    freq_y=[vect]; freq_y=freq_y(1:freqmaxima+1);
    datos_plot_freqc{k}{2}=freq_y;
    datos_plot_freqc{k}{1}=freq_x;
elseif freq_dist(1)==4.2
    %Gráfico de Poisson ZM de la clase (a,b,1)
    tabla1{k,3}='Poisson ZM (lambda)';
    freqmaxima=max(datos_plot_freqd{k}{1});
    freq_x=0:freqmaxima;
    freq_y=zero_mod_nbin(freq_dist(2),c_ZM,freqmaxima);
    freq_y=freq_y(1:freqmaxima+1);
    datos_plot_freqc{k}{2}=freq_y;
    datos_plot_freqc{k}{1}=freq_x;
elseif freq_dist(1)==4.3

```

```

%Gráfico de Poisson ZT de la clase (a,b,1)
    tablal{k,3}='Poisson ZT (lambda)';
    freqmaxima=max(datos_plot_freqd{k}{1});
    freq_x=0:freqmaxima;
    freq_y=zero_mod_nbin(freq_dist(2),c_ZM,freqmaxima);
    freq_y=freq_y(1:freqmaxima+1);
    datos_plot_freqc{k}{2}=freq_y;
    datos_plot_freqc{k}{1}=freq_x;
end

tablal{k,4}=freq_dist(2);
tablal{k,5}=freq_dist(3);

%Codificación de las distribuciones para poder extraer la que mejor
%características presenta para la distribución estudiada:

    %5.Exponencial          8.Inv_Gaussiana          11.LogNormal
    %6.Gamma                9.Logistic              12.Normal
    %7.Gen Pareto           10.LogLogistic         13.Weibull

Mat_Fit_Test_sev=genvarname(['Mat_Fit_Test_sev_cluster_',num2str(k)],who);
mat_resumen_sev=[longit 0 0 0 0 0;
    5 submatteest_exp_sev, round(pV_exp_sev,2),
    round(param_mu_exp_sev,2), 0, 0;
    6 submatteest_gamma_sev, round(pV_gamma_sev,2),
    round(param_a_gamma_sev,2), round(param_b_gamma_sev,2), 0;
    7 submatteest_gpareto_sev, pV_gpareto_sev,
    round(param_k_gpareto_sev,2) round(param_sigma_gpareto_sev,2)
    round(param_theta_gpareto_sev,2);
    8 submatteest_InvGauss_sev round(pV_InvGauss_sev,2)
    round(param_mu_InvGauss_sev,2) round(param_lambda_InvGauss_sev,2) 0;

    9 submatteest_Logi_sev round(pV_Logi_sev,2)
    round(param_mu_Logi_sev,2) round(param_sigma_Logi_sev,2) 0;
    10 submatteest_LogLogi_sev round(pV_LogLogi_sev,2)
    round(param_mu_LogLogi_sev,2) round(param_sigma_LogLogi_sev,2) 0;
    11 submatteest_LogNorm_sev round(pV_LogNorm_sev,2)
    round(param_mu_LogNorm_sev,2) round(param_sigma_LogNorm_sev,2) 0;
    12 submatteest_Norm_sev round(pV_Norm_sev,2)
    round(param_mu_Norm_sev,2) round(param_sigma_Norm_sev,2) 0;
    13 submatteest_Weibull_sev round(pV_Weibull_sev,2)
    round(param_A_Weibull_sev,2) round(param_B_Weibull_sev,2) 0];
format long g
    eval([Mat_Fit_Test_sev '=mat_resumen_sev'])

Sev_dist_cluster=genvarname(['Sev_dist_cluster_',num2str(k)],who);
max_pv_sev=max(mat_resumen_sev(2:10,3)); [row2
col2]=find(mat_resumen_sev==max_pv_sev);
[sev_dist]=mat_resumen_sev(row2,[1 4 5 6]);
eval([Sev_dist_cluster '=[sev_dist]']);

```




```

datos_plot_sevd{k}=cell(2);
datos_plot_sevc{k}{1}=dist_smooth(:,1)';
datos_plot_sevc{k}{2}=dist_smooth(:,2)'./(h_span/20);
datos_plot_sevd{k}{1}=(disc_x+(h_span/2))';
disc_y=zeros(1,disc_x_1); num_no_nan=0;
for i=1:length(submatrix3)
    idx=floor(submatrix3(i)/h_span)+1;
    if isnan(idx)
    else
        disc_y(idx)=disc_y(idx) + 1;
        num_no_nan=num_no_nan+1;
    end
end
disc_y = disc_y./num_no_nan./h_span;
datos_plot_sevd{k}{2}=disc_y';

%Aplicación del Método recursivo de Panjer para el cálculo de la
%distribución de pérdidas agregadas
dist_agloss_panjer=genvarname(['dist_agloss_cluster_',num2str(k)],who);
    %Aplicación del Método Recursivo de Panjer
    dist_agloss_p=panjer(freq_dist,c_ZM,dist_discret,max_loss_disc,h_span);
    max_loss=max(dist_agloss_p(:,1));
    distr_agloss=[dist_agloss_p(:,1);[max_loss+h_span]];
    %Construcción de la función de distribución de pérdidas
    dist_funct_agloss=cumsum(dist_agloss_p(:,2));

    dist_funct_agloss=[dist_funct_agloss; 1];
    %Matriz para reflejar las pérdidas, su probabilidad y probabilidad
    acumulada
    dist_agloss=[distr_agloss dist_funct_agloss];
    VaR_pos=percentil/100;
    VaR=posit_distrib(dist_agloss(:,2),VaR_pos);
    [VaR_clusters(k)]=[dist_agloss(VaR,1)];
eval(['dist_agloss_panjer '=dist_agloss']);
tabla2{k,4}=[dist_agloss(VaR,1)];

datos_grafico{1}=[datos_grafico{1};dist_agloss_p(:,1)'];
datos_grafico{2}=[datos_grafico{2};dist_agloss_p(:,2)'];

%Información de cada cluster para la posterior tarificación
cluster_info=genvarname(['info_cluster_',num2str(k)],who);
    mean_cluster=sum([dist_agloss_p(:,1)].*[dist_agloss_p(:,2)]);
    [mean_clusters(k)]=mean_cluster*(1+perc_costs);
    std_cluster=sqrt((sum((dist_agloss_p(:,1)-mean_cluster).^2)/longit);
    [std_clusters(k)]=std_cluster;
    [info]=[longit mean_cluster std_cluster];
eval(['cluster_info '=[info]']);
[exclaims(k)]=mean_clusters(k)*ncc(k);
total_exclaims=round(sum(exclaims),2);
tabla2{k,2}=round(mean_cluster,2);tabla3{k,3}=round(mean_cluster,2);
tabla2{k,3}=round(std_cluster,2);

```

```

%
%
%
%-----
%
for h=1:simul;
    for j=1:longit;
        rand_prob=rand();
        [pos]=[posit_distrib(dist_funct_agloss,rand_prob)];
        [claim(j)]=round([dist_agloss(pos,1)*(1+perc_costs)],2);
        claim=sum(claim);
    end
    [claims_total(h)]=claim;
end
[claims_dist]=[sort(claims_total)];
[reserv(k)]=round(prctile(claims_dist,percentil),2);
tot_reserv=sum(reserv);
tabla3{k,6}=reserv(k);
[n_clients(k)]=longit;
End

%CÁLCULOS PARA LA TARIFICACIÓN DE SEGUROS POR CLÚSTER
%-----
for a=1:m %Bucle para la tarificación de seguros
    prem_cluster=genvarname(['premier_cluster_',num2str(a)],who);
    [premium]=[round(mean_clusters(a)*(1+k1)+k2*VaR_clusters(a),2)];
    [tot_prem(a)]=[n_clients(a)*premium]; [tot_premiums]=sum(tot_prem);
    [ratio_prem_res]=[n_clients(a)*premium/reserv(a)];
    eval(['premier_cluster ','=[premium ratio_prem_res]']);
    tabla3{a,4}=premium;
    tabla3{a,5}=[n_clients(a)]*premium;
    tabla3{a,7}=round(ratio_prem_res,4);
    tabla4{1,1}=sum(ncc);
    tabla4{1,2}=total_explclaims;
    tabla4{1,3}=tot_premiums;
    tabla4{1,4}=round(tot_premiums/total_explclaims,4);
    tabla4{1,5}=tot_reserv;
    tabla4{1,6}=round(tot_premiums/tot_reserv,4);
end
%Representación gráfica de algoritmo de Panjer.
cols_x=datos_grafico{1}';
cols_y=datos_grafico{2}';
datos_grafico{1}=[];
datos_grafico{2}=[];
[r c]=size(cols_y);
prob_acum=zeros(c,1);
for i=1:r;
    cond_salida=true;
    for j=1:c;
        prob_acum(j)=prob_acum(j)+cols_y(i,j);

        if (prob_acum(j)<0.9975);

```



```
        cond_salida=false;
    end
end

if (cond_salida);
    break
end
datos_grafico{1}=[datos_grafico{1};cols_x(i,:)];
datos_grafico{2}=[datos_grafico{2};cols_y(i,:)];
end
ret{1}=tabla1;
ret{2}=datos_grafico;
ret{3}=tabla2;
ret{4}=datos_plot_freqd; ret{7}=datos_plot_freqc;
ret{5}=datos_plot_sevc; ret{6}=datos_plot_sevd;
ret{8}=tabla3;
ret{9}=tabla4;

end
```

Anexo II. Código utilizado para la realización del Test Chi-Cuadrado para las variables discretas

```

%TEST CHI-CUADRADO. Esta función hace el contraste no paramétrico que
%comprueba si una distribución empírica se ajusta a una distribución
%teórica, siendo el output de la función el resultado del test (0 si se
%acepta la hipótesis nula y 1 si se rechaza) y el correspondiente p-Valor.

function [res_test pV]=chisqr_test(v, p, alpha)
%v: Es el vector que recoge las observaciones, datos empíricos
%p: Es el vector correspondiente a los valores teóricos
%alpha: Nivel de significación, entre 0 y 1, al cual se efectúa el test

ret_idx=1;
ret=zeros(2,length(v));
%Acumula valores observados hacia la derecha
for orig_idx=1:length(v)
    ret(1,ret_idx)=ret(1,ret_idx)+v(orig_idx);
    ret(2,ret_idx)=ret(2,ret_idx)+p(orig_idx);
    if ret(1,ret_idx)>=5
        ret_idx=min(ret_idx+1 ,length(v));
    end
end
%Acumula la cola de los valores observados
while ret(1,ret_idx)<5
    ret_idx=ret_idx-1;
    ret(1,ret_idx)=ret(1,ret_idx)+ret(1,ret_idx+1);
    ret(2,ret_idx)=ret(2,ret_idx)+ret(2,ret_idx+1);
end
%Acorta el resultado a la dimensión real final
ret=ret(1:2,1:ret_idx);
result=ret;
%Cálculo del estadístico Chi-Cuadrado
chi2_value=sum((result(1,:)-result(2,:)).^2./result(2,:));
chi2_table_value=chi2inv((1-alpha),length(result(1,:)-1));
%Cálculo del p-Valor
pV=round(1-chi2cdf(chi2_value,length(result(1,:)-1),4);

%Resultado del Test Chi-Cuadrado
if chi2_value<chi2_table_value
    res_test=0;
else if chi2_value>chi2_table_value
    res_test=1;
end

end

```



Anexo III. Código para cálculo de la distribución Binomial ZM

```
%DISTRIBUCIÓN BINOMIAL ZM. Código para construir la distribución de
%probabilidad Binomial de la clase (a,b,1) incluyendo una modificación de la
%probabilidad en el punto 0 de la función.

function [v_ZM]=zero_mod_bino(n,p,zm_p)
%Proyección de la distribución teórica sin modificar
for i=1:n+1
    [v_prob(i)]=binopdf(i-1,n,p);
end
%Cálculo del parámetro c_ZM para corregir las probabilidades siguientes a 0
    c_ZM=(1-zm_p)/(1-v_prob(1));

%Cálculo de la distribución Binomial ZM
for i=1:n+1
    if i==1
        [v_prob_mod(i)]=zm_p;
    else if i ~= 1
        [v_prob_mod(i)]=(v_prob(i))*c_ZM;
    end
end
[v_ZM]=[v_prob_mod];
format short
end
end
```

Anexo IV. Código para cálculo de la distribución Binomial ZT

```
%DISTRIBUCIÓN BINOMIAL ZT. Código para construir la distribución de
%probabilidad Binomial de la clase (a,b,1) truncando la probabilidad en el
%punto 0 de la función.

function [v_ZT]=zero_trun_bino(n,p)
%Proyección de la distribución teórica sin truncar
for i=1:n+1
    [v_prob(i)]=binopdf(i-1,n,p);
end
%Cálculo del parámetro c_ZT para corregir las probabilidades siguientes a 0
    c_ZT=1/(1-v_prob(1));

%Cálculo de la distribución Binomial ZT
for i=1:n+1
    if i==1
        [v_prob_trun(i)]=0;
    else if i ~= 1
        [v_prob_trun(i)]=(v_prob(i))*c_ZT;
    end
end
[v_ZT]=[v_prob_trun];
format short
end
end
```

Anexo V. Código para cálculo de la distribución Binomial Negativa ZM

```
%DISTRIBUCIÓN BINOMIAL NEGATIVA ZM. Código para construir la distribución
%de probabilidad Binomial Negativa de la clase (a,b,1) incluyendo una
%modificación de la probabilidad en el punto 0 de la función.

function [v_ZM]=zero_mod_nbin(R,P,zm_p,n)
%Proyección de la distribución teórica sin modificar
for i=1:n+1
    [v_prob(i)]=nbinpdf(i-1,R,P);
end
%Cálculo del parámetro c_ZM para corregir las probabilidades siguientes a 0
    c_ZM=(1-zm_p)/(1-v_prob(1));

%Cálculo de la distribución Binomial Negativa ZM
for i=1:n+1
    if i==1
        [v_prob_mod(i)]=zm_p;
    else if i ~= 1
        [v_prob_mod(i)]=(v_prob(i))*c_ZM;
    end
end
[v_ZM]=[v_prob_mod];
format short
end
end
```


Anexo VI. Código para cálculo de la distribución Binomial Negativa ZT

```
%DISTRIBUCIÓN BINOMIAL NEGATIVA ZT. Código para construir la distribución
%de probabilidad Binomial de la clase (a,b,1) truncando la probabilidad en
%el punto 0 de la función.

function [v_ZT]=zero_trun_nbin(R,P,n)
%Proyección de la distribución teórica sin truncar
for i=1:n+1
    [v_prob(i)]=nbinpdf(i-1,R,P);
end
%Cálculo del parámetro c_ZT para corregir las probabilidades siguientes a 0
    c_ZT=1/(1-v_prob(1));

%Cálculo de la distribución Binomial Negativa ZT
for i=1:n+1
    if i==1
        [v_prob_trun(i)]=0;
    else if i ~= 1
        [v_prob_trun(i)]=(v_prob(i))*c_ZT;
    end
end
[v_ZT]=[v_prob_trun];
format short
end
end
```

Anexo VII. Código para cálculo de la distribución Geométrica ZM

```
%DISTRIBUCIÓN GEOMÉTRICA ZM. Código para construir la distribución de
%probabilidad Binomial de la clase (a,b,1) incluyendo una modificación de la
%probabilidad en el punto 0 de la función.

function [v_ZM]=zero_mod_geom(p,zm_p,n)
%Proyección de la distribución teórica sin modificar
for i=1:n+1
    [v_prob(i)]=geopdf(i-1,p);
end
%Cálculo del parámetro c_ZM para corregir las probabilidades siguientes a 0
    c_ZM=(1-zm_p)/(1-v_prob(1));

%Cálculo de la distribución Geométrica ZM
for i=1:n+1
    if i==1
        [v_prob_mod(i)]=zm_p;
    else if i ~= 1
        [v_prob_mod(i)]=(v_prob(i))*c_ZM;
    end
end
[v_ZM]=[v_prob_mod];
format short
end
end
```

Anexo VIII. Código para cálculo de la distribución Geométrica ZT

```
%DISTRIBUCIÓN GEOMÉTRICA ZT. Código para construir la distribución
%de probabilidad Binomial de la clase (a,b,1) truncando la probabilidad en
%el punto 0 de la función.

function [v_ZT]=zero_trun_geom(p,n)
%Proyección de la distribución teórica sin truncar
for i=1:n+1
    [v_prob(i)]=geopdf(i-1,p);
end
%Cálculo del parámetro c_ZT para corregir las probabilidades siguientes a 0
    c_ZT=1/(1-v_prob(1));

%Cálculo de la distribución Geométrica ZT
for i=1:n+1
    if i==1
        [v_prob_trun(i)]=0;
    else if i ~= 1
        [v_prob_trun(i)]=(v_prob(i))*c_ZT;
    end
end
[v_ZT]=[v_prob_trun];
format short
end
end
```

Anexo IX. Código para cálculo de la distribución Poisson ZM

```
%DISTRIBUCIÓN POISSON ZM. Código para construir la distribución de
%probabilidad Binomial de la clase (a,b,1) incluyendo una modificación de la
%probabilidad en el punto 0 de la función.

function [v_ZM]=zero_mod_poiss(lambda,zm_p,n)
%Proyección de la distribución teórica sin modificar
for i=1:n+1
    [v_prob(i)]=poisspdf(i-1,lambda);
end
%Cálculo del parámetro c_ZM para corregir las probabilidades siguientes a 0
    c_ZM=(1-zm_p)/(1-v_prob(1));

%Cálculo de la distribución Poisson ZM
for i=1:n+1
    if i==1
        [v_prob_mod(i)]=zm_p;
    else if i ~= 1
        [v_prob_mod(i)]=(v_prob(i))*c_ZM;
    end
end
[v_ZM]=[v_prob_mod];
format short
end
end
```



Anexo X. Código para cálculo de la distribución Poisson ZT

```
%DISTRIBUCIÓN POISSON ZT. Código para construir la distribución de
%probabilidad Binomial de la clase (a,b,1) truncando la probabilidad en el
%punto 0 de la función.

function [v_ZT]=zero_trun_poiss(lambda,n)
%Proyección de la distribución teórica sin truncar
for i=1:n+1
    [v_prob(i)]=poisspdf(i-1,lambda);
end
%Cálculo del parámetro c_ZT para corregir las probabilidades siguientes a 0
    c_ZT=1/(1-v_prob(1));

%Cálculo de la distribución Binomial ZT
for i=1:n+1
    if i==1
        [v_prob_trun(i)]=0;
    else if i ~= 1
        [v_prob_trun(i)]=(v_prob(i))*c_ZT;
    end
end
[v_ZT]=[v_prob_trun];
format short
end
end
```



Anexo XI. Código para discretización de distribuciones continuas mediante método *Rounding*

```

%MÉTODO ROUNDING DE DISCRETIZACIÓN. Aplicación del método Rounding para
%obtener la versión discretizada de la distribución de severidad continua
%necesaria para poder aplicar el método de Panjer

function [f]=round_discret(dist,param1,param2,param3,h,t)
%"dist". Hace referencia a la codificación de las distribuciones que se ha
%utilizado para el presente aplicativo.
%"param". Según la distribución que más se ajuste a la severidad estudiada, los
%parámetros a introducir como inputs variarán.
%"h". Representa el span que se identifica como un múltiplo de la unidad
%monetaria de tal forma que represente determinadas cuantías en clases.
%"t". Punto hasta el cual se van a realizar la discretización.

%5. Distribución EXPONENCIAL
%-----
if dist==5
    param1=(1/param1);
for j=0:t
    if j==0
        [v_prob_disc(1)]=cdf('Exponential',(h/2),param1);
        [v_mult(1)]=0;
    else if j~=0
        [v_prob_disc(j+1)]=cdf('Exponential',(j*h+h/2),param1)-cdf('Exponential',(j*h-
h/2),param1);
        [v_mult(j+1)]=j*h;
    end
end
    [f]=[v_mult; v_prob_disc]';
end
end

%6. Distribución GAMMA
%-----
if dist==6
for j=0:t
    if j==0
        [v_prob_disc(1)]=cdf('Gamma',(h/2),param1,param2);
        [v_mult(1)]=0;
    else if j~=0
        [v_prob_disc(j+1)]=cdf('Gamma',(j*h+h/2),param1,param2)-cdf('Gamma',(j*h-
h/2),param1,param2);
        [v_mult(j+1)]=j*h;
    end
end
    [f]=[v_mult; v_prob_disc]';

end
    
```



```

end
%7. Distribución GEN PARETO
%-----
if dist==7
for j=0:t
    if j==0
        [v_prob_disc(1)]=cdf('Generalized Pareto', (h/2), param1, param2, param3);
        [v_mult(1)]=0;
    else if j~=0
        [v_prob_disc(j+1)]=cdf('Generalized Pareto', (j*h+h/2), param1, param2, param3) -
cdf('Generalized Pareto', (j*h-h/2), param1, param2, param3);
        [v_mult(j+1)]=j*h;
    end
end
    [f]=[v_mult; v_prob_disc]';
end
end

%8. Distribución INV GAUSSIANA
%-----
if dist==8
for j=0:t
    if j==0
        [v_prob_disc(1)]=cdf('InverseGaussian', (h/2), param1, param2);
        [v_mult(1)]=0;
    else if j~=0
        [v_prob_disc(j+1)]=cdf('InverseGaussian', (j*h+h/2), param1, param2) -
cdf('InverseGaussian', (j*h-h/2), param1, param2);
        [v_mult(j+1)]=j*h;
    end
end
    [f]=[v_mult; v_prob_disc]';
end
end

%9. Distribución LOGISTIC
%-----
if dist==9
for j=0:t
    if j==0
        [v_prob_disc(1)]=cdf('Logistic', (h/2), param1, param2);
        [v_mult(1)]=0;
    else if j~=0
        [v_prob_disc(j+1)]=cdf('Logistic', (j*h+h/2), param1, param2) -cdf('Logistic', (j*h-
h/2), param1, param2);
        [v_mult(j+1)]=j*h;
    end
end
    [f]=[v_mult; v_prob_disc]';
end
end

```

```

%10. Distribución LOG-LOGISTIC
%-----
if dist==10
for j=0:t
    if j==0
        [v_prob_disc(1)]=cdf('LogLogistic', (h/2),param1,param2);
        [v_mult(1)]=0;
    else if j~=0
        [v_prob_disc(j+1)]=cdf('LogLogistic', (j*h+h/2),param1,param2)-
cdf('LogLogistic', (j*h-h/2),param1,param2);
        [v_mult(j+1)]=j*h;
    end
end
    [f]=[v_mult; v_prob_disc]';
end
end

%11. Distribución LOG-NORMAL
%-----
if dist==11
for j=0:t
    if j==0
        [v_prob_disc(1)]=cdf('Lognormal', (h/2),param1,param2);
        [v_mult(1)]=0;
    else if j~=0
        [v_prob_disc(j+1)]=cdf('Lognormal', (j*h+h/2),param1,param2)-cdf('Lognormal', (j*h-
h/2),param1,param2);
        [v_mult(j+1)]=j*h;
    end
end
    [f]=[v_mult; v_prob_disc]';
end
end

%12. Distribución NORMAL
%-----
if dist==12
for j=0:t
    if j==0
        [v_prob_disc(1)]=cdf('Normal', (h/2),param1,param2);
        [v_mult(1)]=0;

    else if j~=0
        [v_prob_disc(j+1)]=cdf('Normal', (j*h+h/2),param1,param2)-cdf('Normal', (j*h-
h/2),param1,param2);
        [v_mult(j+1)]=j*h;
    end
end
    [f]=[v_mult; v_prob_disc]';
end
end

```




```

end

%13. Distribución WEIBULL
%-----
if dist==13
for j=0:t
    if j==0
        [v_prob_disc(1)]=cdf('Weibull',(h/2),param1,param2);
        [v_mult(1)]=0;
    else if j~=0
        [v_prob_disc(j+1)]=cdf('Weibull',(j*h+h/2),param1,param2)-cdf('Weibull',(j*h-
h/2),param1,param2);
        [v_mult(j+1)]=j*h;
    end
end
    [f]=[v_mult; v_prob_disc]';
end
end
end

```

Anexo XII. Código para ejecución del algoritmo recursivo de Panjer

```

%MÉTODO RECURSIVO DE PANJER. Algoritmo que permite calcular, como
%alternativa al método de las convoluciones, la distribución de pérdidas
%agregadas conociendo la distribución que sigue la frecuencia y la
%severidad de los siniestros.

function [x]=panjer(freq_dist,c_ZM,sev_discret,max_loss_disc,span_h)
%freq_dist: Vector que incluye el identificador de la distribución y los
%parámetros que la definen
dist=freq_dist(1);
param1=freq_dist(2); param2=freq_dist(3);
%c_ZM: Modificación de probabilidad en 0 para las distribuciones de la
%clase (a,b,1).
%sev_discret: Versión discretizada de la distribución de severidad.
%max_loss: Máxima pérdida que puede llegar a suceder.
%span_h: Medida utilizada para la discretización de la severidad.

%Caso de la Distribución BINOMIAL
%-----
if dist==1.1
    param_n_bin=param1; param_p_bin=param2;
    a=-((param_p_bin)/(1-param_p_bin));
    b=(param_p_bin*(param_n_bin+1))/(1-param_p_bin);
    fx(1)=sev_discret(1,2); %MATLAB no tiene posición cero, por lo que este comando es
    la probabilidad fx(0) y luego se ajustan las posiciones
    fs(1)=(1+param_p_bin*(fx(1)-1))^(param_n_bin);

%Caso de la Distriución BINOMIAL ZM y BINOMIAL ZT
%-----
else if dist==1.2|dist==1.3;
    param_n_bin=param1; param_p_bin=param2;
    a=-((param_p_bin)/(1-param_p_bin));
    b=(param_p_bin*(param_n_bin+1))/(1-param_p_bin);
    fx(1)=sev_discret(1,2);
    if dist==1.2
        p(1)=c_ZM;
        p(2)=(1-c_ZM)*(param_n_bin*(1-param_p_bin)^(param_n_bin-1)*param_p_bin)/(1-(1-
        param_p_bin)^param_n_bin);
        fs(1)=c_ZM+(1-c_ZM)*(((1+param_p_bin*(fx(1)-1))^(param_n_bin))- (1-
        param_p_bin)^param_n_bin)/(1-(1-param_p_bin)^param_n_bin);
    else dist==1.3
        p(1)=0;
        p(2)=(param_n_bin*((1-param_p_bin)^(param_n_bin-1))*param_p_bin)/(1-(1-
        param_p_bin)^param_n_bin);
        fs(1)=(((1+param_p_bin*(fx(1)-1))^(param_n_bin))- (1-
        param_p_bin)^param_n_bin)/(1-(1-param_p_bin)^param_n_bin);
    end
    [p]=[p(1) p(2)];
end

```

```

end
%Caso de la Distribución BINOMIAL NEGATIVA
%-----
if dist==2.1
    param_r_nbin=param1; param_p_nbin=param2; beta=(1-param_p_nbin)/param_p_nbin;
    a=beta/(1+beta); b=((param_r_nbin-1)*beta)/(1+beta);
    fx(1)=sev_discret(1,2);
    fs(1)=(1+beta*(1-fx(1)))^(-param_r_nbin);

%Caso de la Distribución BINOMIAL NEGATIVA ZM Y ZT
%-----
else if dist==2.2|dist==2.3
    param_r_nbin=param1; param_p_nbin=param2; beta=(1-param_p_nbin)/param_p_nbin;
    a=beta/(1+beta); b=((param_r_nbin-1)*beta)/(1+beta);
    fx(1)=sev_discret(1,2);
    if dist==2.2
        p(1)=c_ZM;
        p(2)=(1-c_ZM)*(param_r_nbin*beta)/(((1+beta)^(param_r_nbin+1))-(1+beta));
        fs(1)=c_ZM+(1-c_ZM)*(((1+beta*(1-fx(1)))^(-param_r_nbin))-(1+beta)^(-
param_r_nbin))/(1-(1+beta)^(-param_r_nbin));
    else dist==2.3
        p(1)=0;
        p(2)=(param_r_nbin*beta)/(((1+beta)^(param_r_nbin+1))-(1+beta));
        fs(1)=(((1+beta*(1-fx(1)))^(-param_r_nbin))-(1+beta)^(-param_r_nbin))/(1-
(1+beta)^(-param_r_nbin));
    end
    [p]=[p(1) p(2)];
end
end
%Caso de la Distribución GEOMÉTRICA
%-----
if dist==3.1
    param_p_geo=param1; beta=(1-param_p_geo)/param_p_geo;
    a=beta/(1+beta); b=0;
    fx(1)=sev_discret(1,2);
    fs(1)=(1+beta*(1-fx(1)))^(-1);

%Caso de la Distribución GEOMÉTRICA ZM y ZT
%-----
else if dist==3.2|dist==3.3
    param_p_geo=param1; beta=(1-param_p_geo)/param_p_geo;
    a=beta/(1+beta); b=0;
    fx(1)=sev_discret(1,2);
    if dist==3.2
        p(1)=c_ZM;
        p(2)=(1-c_ZM)*1/(1+beta);
        fs(1)=c_ZM+(1-c_ZM)*(fx(1)/(1+beta*(1-fx(1))));
    else dist==3.3
        p(1)=0;

        p(2)=1/(1+beta);
    end
end

```



```

        fs(1)=(fx(1)/(1+beta*(1-fx(1))));
    end
    [p]=[p(1) p(2)];
    end
end
%Caso de la Distribución POISSON
%-----
if dist==4.1
    param_lambda_poisson=param1;
    a=0;b=param_lambda_poisson;
    fx(1)=sev_discret(1,2);
    fs(1)=exp(-param_lambda_poisson*(1-fx(1)));

%Caso de la Distribución POISSON ZM y ZT
%-----
else if dist==4.2|dist==4.3
    param_lambda_poisson=param1;
    a=0;b=param_lambda_poisson;
    fx(1)=sev_discret(1,2);
    if dist==4.2
        p(1)=c_ZM;
        p(2)=(1-c_ZM)*param_lambda_poisson/((exp(param_lambda_poisson))-1);
        fs(1)=c_ZM+(1-c_ZM)*(exp(param_lambda_poisson*fx(1))-
1)/(exp(param_lambda_poisson)-1);
    else %dist==4.3
        p(1)=0;
        p(2)=param_lambda_poisson/((exp(param_lambda_poisson))-1);
        fs(1)=(exp(param_lambda_poisson*fx(1))-1)/(exp(param_lambda_poisson)-1);
    end
    [p]=[p(1) p(2)];
    end
end
%
%-----
%Fórmula recursiva de Panjer para los casos de distribuciones de frecuencias de la
clase (a,b,0)
if dist==[1.1]|dist==[2.1]|dist==[3.1]|dist==[4.1]
d=(1-a*sev_discret(1,2));%d --> Denominador común de la fórmula recursiva de Panjer

for j=1:(100000/span_h);%Se fija realizar el cálculo de la fórmula recursiva de Panjer
100 veces para asegurar
    %que abarca el mayor rango de probabilidades
if j==1;
    [vector(j)]=0;
else if j>1;
    [vector(j)]=(j-1)*span_h;
end
    vec=[vector(j)]; [vec]=vec';
end
end
for x=2:100000/span_h
    m_lim=max_loss_disc;

```



```

[r c]=size(sev_discret(:,1)); x_lim=r;
limit=min(x_lim,m_lim);

if x<=limit;
    for j=1:(x-1);
        [fst(j)]=sum((a+b*j/(x-1))*sev_discret(j+1,2)*fs(x-j));
        [fs(x)]=sum(fst)/d;
    end

elseif x>limit;
    for j=1:(limit-1);
        [fst(j)]=sum((a+b*j/(x-1))*sev_discret(j+1,2)*fs(x-j));
        [fs(x)]=sum(fst)/d;
    end
    [fs(x)]=[abs(fs(x))];
    x=[fs(x)];
end
end
x=fs'; x=[vector; fs]; x=x'; x=round(x,6);

%
%Fórmula recursiva de Panjer para los casos de distribuciones de frecuencias de la
clase (a,b,1)

else
dist==[1.2]|dist==[1.3]|dist==[2.2]|dist==[2.3]|dist==[3.2]|dist==[3.3]|dist==[4.2]|di
st==[4.3]
init=(p(2)-(a+b)*p(1));%--> Utilizado para simplificación de la fórmula
d=(1-a*sev_discret(1,2));%d--> Denominador común de la fórmula recursiva de Panjer

for j=1:100000/span_h%Se fija realizar el cálculo de la fórmula recursiva de Panjer
100 veces para asegurar
        %que abarca el mayor rango de probabilidades
if j==1;
    [vector(j)]=0;
else if j>1;
    [vector(j)]=(j-1)*span_h;
end
vec=[vector(j)]; [vec]=vec';
end
end

for x=2:100000/span_h
    m_lim=max_loss_disc;
    [r c]=size(sev_discret(:,1)); x_lim=r;
    limit=min(x_lim,m_lim);

    if x<=limit;
        for j=1:(x-1);
            [fst(j)]=sum((a+b*j/(x-1))*sev_discret(j+1,2)*fs(x-j));

```



```
[fs(x)]=(init*sev_discret(x,2)+sum(fst))/d;
end

elseif x>limit;
for j=1:(limit-1);
[fst(j)]=sum((a+b*j/(x-1))*sev_discret(j+1,2)*fs(x-j));
[fs(x)]=sum(fst)/d;
end
[fs(x)]=[abs(fs(x))];
x=[fs(x)];
end
end
x=fs'; x=[vector; fs]; x=x'; x=round(x,6);
end
end
```



Anexo XIII. Código para seleccionar el Value-at-Risk y severidades simuladas

```
%POSICIÓN DE VECTOR EN RELACIÓN A UN CALOR DETERMINADO. Esta función se  
%utiliza para el cálculo de reservas y el VaR. Mediante iteraciones, se  
%localiza la posición del vector y valor de la cuantía de severidad según  
%el número aleatorio que surja de la simulación o el percentil definido.
```

```
function [pos]=posit_distrib(dist_func, number)
```

```
%dist_func: Es la matriz que contiene los vectores de severidad y su  
%correspondiente función de distribución.
```

```
%number: es el número aleatorio simulado en el caso de las reservas o el  
%percentil que se busca de la distribución, según la aplicación.
```

```
    for i=1:length(dist_func);  
        if dist_func(i)>number;  
            [pos]=i;  
            return  
        [pos]=[pos];  
    end  
end  
end
```



Anexo XIV. Código de interfaz gráfica

```

classdef App1 < matlab.apps.AppBase
    % Properties that correspond to app components
    properties (Access = public)
        TarificadordeSegurosNoVidaUIFigure matlab.ui.Figure
        AbrirarchivoButton matlab.ui.control.Button
        BasedeDatosaanalizarLabel matlab.ui.control.Label
        TextArchivoEntrada matlab.ui.control.EditField
        NiveldeSignificacinLabel matlab.ui.control.Label
        NumNivelSignif matlab.ui.control.NumericEditField
        AjustesyTarificacinButton matlab.ui.control.Button
        NmerodeclsteresLabel matlab.ui.control.Label
        Numerodeclusters matlab.ui.control.NumericEditField
        INPUTSPARAANLISISDECLSTERSLabel matlab.ui.control.Label
        INPUTSPARATARIFICACINDESEGUROSLabel matlab.ui.control.Label
        Modificacindeprobabilidadeden0EditFieldLabel matlab.ui.control.Label
        c_ZMMod matlab.ui.control.NumericEditField
        RecargodeseguridadEditFieldLabel matlab.ui.control.Label
        k1_seg matlab.ui.control.NumericEditField
        RecargoderiesgoEditFieldLabel matlab.ui.control.Label
        k2_riesgo matlab.ui.control.NumericEditField
        PercentildereservasyVaRLabel matlab.ui.control.Label
        perc_res matlab.ui.control.NumericEditField
        TabGroup matlab.ui.container.TabGroup
        AjustedeDistribucionesTab matlab.ui.container.Tab
        UITable matlab.ui.control.Table
        DistribucinPrdidadasAgregadasTab matlab.ui.container.Tab
        UITable2 matlab.ui.control.Table
        GrficosPrdidadasAgregadasProductoButton matlab.ui.control.Button
        GrficodeclsterEditField_2Label matlab.ui.control.Label
        GrficodeclsterEditField_2 matlab.ui.control.NumericEditField
        GrficosPrdidadasAgregadasporClsterButton matlab.ui.control.Button
        TarificacindeSegurosTab matlab.ui.container.Tab
        UITable3 matlab.ui.control.Table
        ResumenCarteraProductoTab matlab.ui.container.Tab
        UITable4 matlab.ui.control.Table
        TextArea matlab.ui.control.TextArea
        GrficodeclsterEditFieldLabel matlab.ui.control.Label
        GrficodeclsterEditField matlab.ui.control.NumericEditField
        GrficodeFrecuenciaButton matlab.ui.control.Button
        GrficodeSeveridadButton matlab.ui.control.Button
        NmerodesimulacionesdesiniestralidadEditFieldLabel matlab.ui.control.Label
        n_simulaciones matlab.ui.control.NumericEditField
        CostesgestindesiniestrosLabel matlab.ui.control.Label
        cost_gest matlab.ui.control.NumericEditField
        SpanparadiscretizacindeseveridadesEditFieldLabel matlab.ui.control.Label
        Span matlab.ui.control.NumericEditField
        Label matlab.ui.control.Label
    end
end

```




```

Label_2                matlab.ui.control.Label
Label_3                matlab.ui.control.Label
Label_4                matlab.ui.control.Label
Label_5                matlab.ui.control.Label
end
properties (Access = private)
    resultadocalculo; % Description
end
methods (Access = private)
    % Button pushed function: AbrirarchivoButton
    function abrirArchivo(app, event)
        [nombre, carpeta]=uigetfile('*.xlsx');
        ruta=fullfile(carpeta, nombre);
        app.TextArchivoEntrada.Value=ruta;
        app.AjustesyTarificacinButton.Enable='on';
    end
    % Button pushed function: AjustesyTarificacinButton
    function CalculaAjuste(app, event)
        alpha=app.NumNivelSignif.Value/100;
        m=app.Numerodeclusters.Value;
        c_ZM=app.c_ZMMod.Value;
        h_span=app.Span.Value;
        simul=app.n_simulaciones.Value;
        perc_costs=app.cost_gest.Value/100;
        percentil=app.perc_res.Value;
        k1=app.k1_seg.Value/100;
        k2=app.k2_riesgo.Value/100;
        path=app.TextArchivoEntrada.Value;

app.resultadocalculo=Tarificador_NoVida(path,m,alpha,c_ZM,h_span,simul,perc_costs,perc
entil,k1,k2);

        datostabla=app.resultadocalculo{1};
        app.UITable.Data=datostabla;
        app.UITable2.Data=app.resultadocalculo{3};
        app.UITable3.Data=app.resultadocalculo{8};
        app.UITable4.Data=app.resultadocalculo{9};

    end
    % Button pushed function:
    % GrficosPrdidasAgregadasProductoButton
    function mostrarGrafico(app, event)
        m=app.Numerodeclusters.Value;
        datosgrafico = app.resultadocalculo{2};
        plot(datosgrafico{1},datosgrafico{2}, 'linewidth', 2);
        legend(compose('Distrib. Pérdidas Agregadas de Panjer - Clúster %d',1:m));
    end
    % Button pushed function: GrficodeFrecuenciaButton
    function PlotFrecuencia(app, event)
        k=app.GrficodeclsterEditField.Value;

        valores_c=app.resultadocalculo{7}{k};

```



```

        valores_d=app.resultadocalculo{4}{k};
        bar(valores_d{1},valores_d{2});
        hold on;
        plot(valores_c{1},valores_c{2},'r','linewidth',2);
        legend('Valores de frecuencias observados', 'Ajuste de distribución
teórica');
        hold off;
        title('Frecuencia cluster '+string(k));
    end
    % Button pushed function: GrficodeSeveridadButton
    function PlotSeveridad(app, event)
        k=app.GrficodeclsterEditField.Value;
        valores_c=app.resultadocalculo{5}{k};
        valores_d=app.resultadocalculo{6}{k};
        bar(valores_d{1},valores_d{2});
        hold on;
        plot(valores_c{1},valores_c{2},'r','linewidth',2);
        legend('Severidad empírica', 'Ajuste de distribución teórica');
        hold off;
        title('Severidad cluster '+string(k));
    end
    % Button pushed function:
    % GrficosPrdidiasAgregadasporClsterButton
    function PerdidasCluster(app, event)
        k=app.GrficodeclsterEditField_2.Value;
        m=app.Numerodeclusters.Value;
        datosgrafico = app.resultadocalculo{2};
        plot(datosgrafico{1}(:,k),datosgrafico{2}(:,k), 'linewidth', 2);
        legend('Distrib. Pérdidas Agregadas de Panjer - Clúster '+string(k));
    end
end
% App initialization and construction
methods (Access = private)
    % Create UIFigure and components
    function createComponents(app)
        % Create TarificadordeSegurosNoVidaUIFigure
        app.TarificadordeSegurosNoVidaUIFigure = uifigure;
        app.TarificadordeSegurosNoVidaUIFigure.Color = [0.2706 0.451 0.5882];
        app.TarificadordeSegurosNoVidaUIFigure.Position = [100 100 767 526];
        app.TarificadordeSegurosNoVidaUIFigure.Name = 'Tarificador de Seguros No
Vida';

        app.TarificadordeSegurosNoVidaUIFigure.Resize = 'off';
        setAutoResize(app, app.TarificadordeSegurosNoVidaUIFigure, true)
        % Create AbrirarchivoButton
        app.AbrirarchivoButton = uibutton(app.TarificadordeSegurosNoVidaUIFigure,
'push');

        app.AbrirarchivoButton.ButtonPushedFcn = createCallbackFcn(app,
@abrirArchivo, true);

        app.AbrirarchivoButton.FontWeight = 'bold';
        app.AbrirarchivoButton.Position = [549 489 100 22];
    end
end

```



```

app.AbrirarchivoButton.Text = 'Abrir archivo';
% Create BasedeDatosaanalizarLabel
app.BasedeDatosaanalizarLabel =
uilabel(app.TarificadordeSegurosNoVidaUIFigure);
app.BasedeDatosaanalizarLabel.HorizontalAlignment = 'right';
app.BasedeDatosaanalizarLabel.FontColor = [1 1 1];
app.BasedeDatosaanalizarLabel.Position = [65 493 142 15];
app.BasedeDatosaanalizarLabel.Text = 'Base de Datos a analizar';
% Create TextArchivoEntrada
app.TextArchivoEntrada =
uieditfield(app.TarificadordeSegurosNoVidaUIFigure, 'text');
app.TextArchivoEntrada.Editable = 'off';
app.TextArchivoEntrada.FontColor = [1 1 1];
app.TextArchivoEntrada.BackgroundColor = [0.9882 0.9686 0.9686];
app.TextArchivoEntrada.Position = [222 489 304 22];
% Create NiveldesignificacinLabel
app.NiveldesignificacinLabel =
uilabel(app.TarificadordeSegurosNoVidaUIFigure);
app.NiveldesignificacinLabel.FontColor = [0.9608 0.9608 0.9608];
app.NiveldesignificacinLabel.Position = [67 389 121 15];
app.NiveldesignificacinLabel.Text = 'Nivel de significación';
% Create NumNivelSignif
app.NumNivelSignif = uieditfield(app.TarificadordeSegurosNoVidaUIFigure,
'numeric');
app.NumNivelSignif.Position = [292 385 25 22];
app.NumNivelSignif.Value = 5;
% Create AjustesyTarificacinButton
app.AjustesyTarificacinButton =
uibutton(app.TarificadordeSegurosNoVidaUIFigure, 'push');
app.AjustesyTarificacinButton.ButtonPushedFcn = createCallbackFcn(app,
@CalculaAjuste, true);
app.AjustesyTarificacinButton.Enable = 'off';
app.AjustesyTarificacinButton.FontWeight = 'bold';
app.AjustesyTarificacinButton.Position = [292 239 187 27];
app.AjustesyTarificacinButton.Text = 'Ajustes y Tarificación';
% Create NmerodeclsteresLabel
app.NmerodeclsteresLabel =
uilabel(app.TarificadordeSegurosNoVidaUIFigure);
app.NmerodeclsteresLabel.FontColor = [1 1 1];
app.NmerodeclsteresLabel.Position = [67 417 118 15];
app.NmerodeclsteresLabel.Text = 'Número de clústeres';
% Create Numerodeclusters
app.Numerodeclusters = uieditfield(app.TarificadordeSegurosNoVidaUIFigure,
'numeric');
app.Numerodeclusters.Position = [293 419 23 22];
app.Numerodeclusters.Value = 1;

% Create INPUTSPARAANLISISDECLSTERESLabel
app.INPUTSPARAANLISISDECLSTERESLabel =
uilabel(app.TarificadordeSegurosNoVidaUIFigure);
app.INPUTSPARAANLISISDECLSTERESLabel.HorizontalAlignment = 'center';

```



```

app.INPUTSPARAANLISISDECLSTERSLabel.VerticalAlignment = 'center';
app.INPUTSPARAANLISISDECLSTERSLabel.FontSize = 14;
app.INPUTSPARAANLISISDECLSTERSLabel.FontWeight = 'bold';
app.INPUTSPARAANLISISDECLSTERSLabel.FontColor = [1 1 1];
app.INPUTSPARAANLISISDECLSTERSLabel.Position = [55.5 447 273 18];
app.INPUTSPARAANLISISDECLSTERSLabel.Text = 'INPUTS PARA ANÁLISIS DE
CLÚSTERS';
% Create INPUTSPARATARIFICACINDESEGUROSLabel
app.INPUTSPARATARIFICACINDESEGUROSLabel =
uilabel(app.TarificadordeSegurosNoVidaUIFigure);
app.INPUTSPARATARIFICACINDESEGUROSLabel.HorizontalAlignment = 'center';
app.INPUTSPARATARIFICACINDESEGUROSLabel.VerticalAlignment = 'center';
app.INPUTSPARATARIFICACINDESEGUROSLabel.FontSize = 14;
app.INPUTSPARATARIFICACINDESEGUROSLabel.FontWeight = 'bold';
app.INPUTSPARATARIFICACINDESEGUROSLabel.FontColor = [1 1 1];
app.INPUTSPARATARIFICACINDESEGUROSLabel.Position = [420 447 304 18];
app.INPUTSPARATARIFICACINDESEGUROSLabel.Text = 'INPUTS PARA TARIFICACIÓN
DE SEGUROS';
% Create Modificacindeprobabilidaden0EditFieldLabel
app.Modificacindeprobabilidaden0EditFieldLabel =
uilabel(app.TarificadordeSegurosNoVidaUIFigure);
app.Modificacindeprobabilidaden0EditFieldLabel.FontColor = [1 1 1];
app.Modificacindeprobabilidaden0EditFieldLabel.Position = [69 356 191 18];
app.Modificacindeprobabilidaden0EditFieldLabel.Text = 'Modificación de
probabilidad en 0';
% Create c_ZMMod
app.c_ZMMod = uieditfield(app.TarificadordeSegurosNoVidaUIFigure,
'numeric');
app.c_ZMMod.Position = [279 354 39 22];
app.c_ZMMod.Value = 0.35;
% Create RecargodeseguridadEditFieldLabel
app.RecargodeseguridadEditFieldLabel =
uilabel(app.TarificadordeSegurosNoVidaUIFigure);
app.RecargodeseguridadEditFieldLabel.FontColor = [1 1 1];
app.RecargodeseguridadEditFieldLabel.Position = [428 316 127 15];
app.RecargodeseguridadEditFieldLabel.Text = 'Recargo de seguridad';
% Create k1_seg
app.k1_seg = uieditfield(app.TarificadordeSegurosNoVidaUIFigure,
'numeric');
app.k1_seg.Limits = [0 Inf];
app.k1_seg.Position = [690 312 44 22];
app.k1_seg.Value = 2.5;
% Create RecargoderiesgoEditFieldLabel
app.RecargoderiesgoEditFieldLabel =
uilabel(app.TarificadordeSegurosNoVidaUIFigure);
app.RecargoderiesgoEditFieldLabel.FontColor = [1 1 1];

app.RecargoderiesgoEditFieldLabel.Position = [428 290 113 15];
app.RecargoderiesgoEditFieldLabel.Text = 'Recargo de riesgo';
% Create k2_riesgo

```



```

        app.k2_riesgo = uieditfield(app.TarificadordeSegurosNoVidaUIFigure,
'numeric');
        app.k2_riesgo.Limits = [0 Inf];
        app.k2_riesgo.Position = [690 286 44 22];
        app.k2_riesgo.Value = 1;
        % Create PercentildereservasyVaRLabel
        app.PercentildereservasyVaRLabel =
uilabel(app.TarificadordeSegurosNoVidaUIFigure);
        app.PercentildereservasyVaRLabel.FontColor = [0.9255 0.9255 0.9255];
        app.PercentildereservasyVaRLabel.Position = [428 342 156 15];
        app.PercentildereservasyVaRLabel.Text = 'Percentil de reservas y VaR';
        % Create perc_res
        app.perc_res = uieditfield(app.TarificadordeSegurosNoVidaUIFigure,
'numeric');
        app.perc_res.Position = [690 338 44 22];
        app.perc_res.Value = 99.5;
        % Create TabGroup
        app.TabGroup = uitabgroup(app.TarificadordeSegurosNoVidaUIFigure);
        app.TabGroup.Position = [9 50 750 181];
        % Create AjustedeDistribucionesTab
        app.AjustedeDistribucionesTab = uitab(app.TabGroup);
        app.AjustedeDistribucionesTab.Title = 'Ajuste de Distribuciones';
        app.AjustedeDistribucionesTab.Units = 'pixels';
        % Create UITable
        app.UITable =uitable(app.AjustedeDistribucionesTab);
        app.UITable.ColumnName = {'Cluster'; 'Polizas'; 'Distribución
Frecuencias'; 'Par. 1'; 'Par. 2'; 'Distribución Severidad'; 'Par. 1'; 'Par. 2'; 'Par.
3'};
        app.UITable.ColumnWidth = {49, 50, 190, 55, 50, 190, 50, 50, 50};
        app.UITable.RowName = {};
        app.UITable.Position = [7 2 736 150];
        app.UITable.ColumnEditable = [true false false false false false false
false false];
        app.UITable.FontWeight = 'bold';
        % Create DistribucinPrdididasAgregadasTab
        app.DistribucinPrdididasAgregadasTab = uitab(app.TabGroup);
        app.DistribucinPrdididasAgregadasTab.Title = 'Distribución Pérdidas
Agregadas';
        app.DistribucinPrdididasAgregadasTab.Units = 'pixels';
        % Create UITable2
        app.UITable2 =uitable(app.DistribucinPrdididasAgregadasTab);
        app.UITable2.ColumnName = {'Clúster'; 'Media de Distribución'; 'Desviación
Típica Distribución'; 'Value-at-Risk'};
        app.UITable2.ColumnWidth = {50, 138, 178, 84};
        app.UITable2.RowName = {};
        app.UITable2.Position = [24 16 452 129];

        app.UITable2.FontWeight = 'bold';
        % Create GrficosPrdididasAgregadasProductoButton
        app.GrficosPrdididasAgregadasProductoButton =
uibutton(app.DistribucinPrdididasAgregadasTab, 'push');

```



```

        app.GrficosPrdidadasAgregadasProductoButton.ButtonPushedFcn =
createCallbackFcn(app, @mostrarGrafico, true);
        app.GrficosPrdidadasAgregadasProductoButton.Position = [501.5 33 240 26];
        app.GrficosPrdidadasAgregadasProductoButton.Text = 'Gráficos Pérdidas
Agregadas Producto';
        % Create GrficodeclsterEditField_2Label
        app.GrficodeclsterEditField_2Label =
uilabel(app.DistribucinPrdidadasAgregadasTab);
        app.GrficodeclsterEditField_2Label.HorizontalAlignment = 'right';
        app.GrficodeclsterEditField_2Label.FontWeight = 'bold';
        app.GrficodeclsterEditField_2Label.Position = [545 119 107 15];
        app.GrficodeclsterEditField_2Label.Text = 'Gráfico de clúster';
        % Create GrficodeclsterEditField_2
        app.GrficodeclsterEditField_2 =
uieditfield(app.DistribucinPrdidadasAgregadasTab, 'numeric');
        app.GrficodeclsterEditField_2.Limits = [1 Inf];
        app.GrficodeclsterEditField_2.FontWeight = 'bold';
        app.GrficodeclsterEditField_2.Position = [667 115 29 22];
        app.GrficodeclsterEditField_2.Value = 1;
        % Create GrficosPrdidadasAgregadasporClsterButton
        app.GrficosPrdidadasAgregadasporClsterButton =
uibutton(app.DistribucinPrdidadasAgregadasTab, 'push');
        app.GrficosPrdidadasAgregadasporClsterButton.ButtonPushedFcn =
createCallbackFcn(app, @PerdidadasCluster, true);
        app.GrficosPrdidadasAgregadasporClsterButton.Position = [501.5 78 240 26];
        app.GrficosPrdidadasAgregadasporClsterButton.Text = 'Gráficos Pérdidas
Agregadas por Clúster';
        % Create TarificacindeSegurosTab
        app.TarificacindeSegurosTab = uitab(app.TabGroup);
        app.TarificacindeSegurosTab.Title = 'Tarificación de Seguros';
        app.TarificacindeSegurosTab.Units = 'pixels';
        % Create UITable3
        app.UITable3 = uitable(app.TarificacindeSegurosTab);
        app.UITable3.ColumnName = {'Clúster'; 'Pólizas'; 'Esperanza siniestro';
'Prima cobrada'; 'Total Primas cobradas'; 'Siniestralidad simulada'; 'Ratio
Primas/siniestralidad'};
        app.UITable3.ColumnWidth = {45, 45, 120, 95, 133, 140, 170};
        app.UITable3.RowName = {};
        app.UITable3.Position = [0 1 750 159];
        app.UITable3.FontWeight = 'bold';
        % Create ResumenCarteraProductoTab
        app.ResumenCarteraProductoTab = uitab(app.TabGroup);
        app.ResumenCarteraProductoTab.Title = 'Resumen Cartera Producto';
        app.ResumenCarteraProductoTab.Units = 'pixels';
        % Create UITable4

        app.UITable4 = uitable(app.ResumenCarteraProductoTab);
        app.UITable4.ColumnName = {'Total Pólizas'; 'Siniest Total Esperada';
'Total Primas cobradas'; 'Primas/Sin. Esperada'; 'Siniestralidad Total simulada';
'Primas/Sin. Simul.'};
        app.UITable4.ColumnWidth = {75, 132, 131, 130, 170, 110};

```



```

app.UITable4.RowName = {};
app.UITable4.Position = [0 44 750 38];
app.UITable4.FontWeight = 'bold';
% Create TextArea
app.TextArea = uitextarea(app.ResumenCarteraProductoTab);
app.TextArea.FontWeight = 'bold';
app.TextArea.BackgroundColor = [0.9373 0.9373 0.9373];
app.TextArea.Position = [17 98 718 35];
app.TextArea.Value = {'A continuación, se muestran los resultados
agregados del proceso de tarificación de la cartera de pólizas para el producto
seleccionado. '};
% Create GrficodeclsterEditFieldLabel
app.GrficodeclsterEditFieldLabel =
uilabel(app.TarificadordeSegurosNoVidaUIFigure);
app.GrficodeclsterEditFieldLabel.HorizontalAlignment = 'right';
app.GrficodeclsterEditFieldLabel.FontWeight = 'bold';
app.GrficodeclsterEditFieldLabel.FontColor = [1 1 1];
app.GrficodeclsterEditFieldLabel.Position = [89 22 107 15];
app.GrficodeclsterEditFieldLabel.Text = 'Gráfico de clúster';
% Create GrficodeclsterEditField
app.GrficodeclsterEditField =
uieditfield(app.TarificadordeSegurosNoVidaUIFigure, 'numeric');
app.GrficodeclsterEditField.Limits = [1 Inf];
app.GrficodeclsterEditField.FontWeight = 'bold';
app.GrficodeclsterEditField.Position = [211 18 29 22];
app.GrficodeclsterEditField.Value = 1;
% Create GrficodeFrecuenciaButton
app.GrficodeFrecuenciaButton =
uibutton(app.TarificadordeSegurosNoVidaUIFigure, 'push');
app.GrficodeFrecuenciaButton.ButtonPushedFcn = createCallbackFcn(app,
@PlotFrecuencia, true);
app.GrficodeFrecuenciaButton.Position = [285 18 200 22];
app.GrficodeFrecuenciaButton.Text = 'Gráfico de Frecuencia';
% Create GrficodeSeveridadButton
app.GrficodeSeveridadButton =
uibutton(app.TarificadordeSegurosNoVidaUIFigure, 'push');
app.GrficodeSeveridadButton.ButtonPushedFcn = createCallbackFcn(app,
@PlotSeveridad, true);
app.GrficodeSeveridadButton.Position = [516 18 200 22];
app.GrficodeSeveridadButton.Text = 'Gráfico de Severidad';
% Create NmerodesimulacionesdesiniestralidadEditFieldLabel
app.NmerodesimulacionesdesiniestralidadEditFieldLabel =
uilabel(app.TarificadordeSegurosNoVidaUIFigure);
app.NmerodesimulacionesdesiniestralidadEditFieldLabel.HorizontalAlignment
= 'right';
app.NmerodesimulacionesdesiniestralidadEditFieldLabel.FontColor = [1 1 1];
app.NmerodesimulacionesdesiniestralidadEditFieldLabel.Position = [428 393
229 15];
app.NmerodesimulacionesdesiniestralidadEditFieldLabel.Text = 'Número de
simulaciones de siniestralidad';
% Create n_simulaciones

```



```

        app.n_simulaciones = uieditfield(app.TarificadordeSegurosNoVidaUIFigure,
'numeric');
        app.n_simulaciones.Limits = [0 Inf];
        app.n_simulaciones.Position = [668 389 66 22];
        app.n_simulaciones.Value = 100;
        % Create CostesgestindesiniestrosLabel
        app.CostesgestindesiniestrosLabel =
uilabel(app.TarificadordeSegurosNoVidaUIFigure);
        app.CostesgestindesiniestrosLabel.FontColor = [0.9686 0.9686 0.9686];
        app.CostesgestindesiniestrosLabel.Position = [428 367 158 15];
        app.CostesgestindesiniestrosLabel.Text = 'Costes gestión de siniestros';
        % Create cost_gest
        app.cost_gest = uieditfield(app.TarificadordeSegurosNoVidaUIFigure,
'numeric');
        app.cost_gest.Position = [690 363 44 22];
        app.cost_gest.Value = 0.75;
        % Create SpanparadiscretizacindeseveridadesEditFieldLabel
        app.SpanparadiscretizacindeseveridadesEditFieldLabel =
uilabel(app.TarificadordeSegurosNoVidaUIFigure);
        app.SpanparadiscretizacindeseveridadesEditFieldLabel.FontColor = [1 1 1];
        app.SpanparadiscretizacindeseveridadesEditFieldLabel.Position = [428 421
229 15];
        app.SpanparadiscretizacindeseveridadesEditFieldLabel.Text = 'Span para
discretización de severidades';
        % Create Span
        app.Span = uieditfield(app.TarificadordeSegurosNoVidaUIFigure, 'numeric');
        app.Span.Limits = [10 5000];
        app.Span.Position = [689 417 44 22];
        app.Span.Value = 50;
        % Create Label
        app.Label = uilabel(app.TarificadordeSegurosNoVidaUIFigure);
        app.Label.FontWeight = 'bold';
        app.Label.FontColor = [1 1 1];
        app.Label.Position = [317 388 25 15];
        app.Label.Text = '%';
        % Create Label_2
        app.Label_2 = uilabel(app.TarificadordeSegurosNoVidaUIFigure);
        app.Label_2.FontWeight = 'bold';
        app.Label_2.FontColor = [1 1 1];
        app.Label_2.Position = [736 366 25 15];
        app.Label_2.Text = '%';
        % Create Label_3
        app.Label_3 = uilabel(app.TarificadordeSegurosNoVidaUIFigure);

        app.Label_3.FontWeight = 'bold';
        app.Label_3.FontColor = [1 1 1];
        app.Label_3.Position = [736 342 25 15];
        app.Label_3.Text = '%';
        % Create Label_4
        app.Label_4 = uilabel(app.TarificadordeSegurosNoVidaUIFigure);
        app.Label_4.FontWeight = 'bold';

```




```

        app.Label_4.FontColor = [1 1 1];
        app.Label_4.Position = [736 315 25 15];
        app.Label_4.Text = '%';
        % Create Label_5
        app.Label_5 = uilabel(app.TarificadordeSegurosNoVidaUIFigure);
        app.Label_5.FontWeight = 'bold';
        app.Label_5.FontColor = [1 1 1];
        app.Label_5.Position = [735 290 25 15];
        app.Label_5.Text = '%';
    end
end
methods (Access = public)
    % Construct app
    function app = App1()
        % Create and configure components
        createComponents(app)
        % Register the app with App Designer
        registerApp(app, app.TarificadordeSegurosNoVidaUIFigure)
        if nargin == 0
            clear app
        end
    end
    % Code that executes before app deletion
    function delete(app)
        % Delete UIFigure when app is deleted
        delete(app.TarificadordeSegurosNoVidaUIFigure)
    end
end
end

```