



BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA

FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS

AJUSTE DE UNA DISTRIBUCIÓN PARETO GENERALIZADA PARA EL ANÁLISIS DE SINIESTRALIDAD

Tesis presentada como requisito para obtener el título de:
Licenciatura en Actuaría

Presenta:
Matthieu Martínez Huerta.

Director de tesis:
Dr. José Raúl Castro Esparza.

Puebla, Pue.
Septiembre 2015

A mi madre y hermano.

Introducción.

Dentro de una entidad aseguradora existen observaciones cuyo comportamiento no se ajusta a los valores esperados de siniestralidad, generando una amenaza para su estabilidad. El comportamiento inusual de una variable aleatoria puede tener más interés que su “normalidad” ampliamente tratada por la teoría clásica del riesgo. La Teoría del Valor Extremo y más concretamente la distribución de Pareto Generalizada permite modelar los siniestros que exceden un determinado umbral para generar una herramienta de análisis de riesgos y prevención.

El presente trabajo de investigación aplica esta teoría, dentro del marco del seguro, teniendo como objetivo fortalecer la siniestralidad esperada minimizando el riesgo que pone en peligro la estabilidad y solvencia de la entidad aseguradora.

Desde el punto de vista de la entidad aseguradora, el principal interés se encuentra en conocer su capacidad financiera para cubrir la siniestralidad esperada, frente a la posibilidad de ceder al reaseguro, siendo el reaseguro de exceso de pérdida, *excess loss*, el que mantiene una relación directa con la distribución de los siniestros que exceden un cierto umbral.

Desde el punto de vista del reasegurador, el interés por modelar los extremos se centra en determinar la esperanza de siniestralidad por encima de la prioridad y la prima pura de dicho reaseguro.

Estimando la distribución del costo de los siniestros y la distribución de frecuencia de siniestralidad, la teoría clásica del riesgo indicaría el importe de reservas necesario, dada una determinada probabilidad de ruina y un horizonte temporal.

Índice general.

Introducción.

1. Antecedentes.	7
1.1 Gestión de riesgo.....	8
1.2 Reaseguro y solvencia.....	11
1.3 Proceso de reservas.....	12
1.4 Reaseguro de exceso de pérdida.....	15
1.5 Limitaciones del Valor Extremo.....	18
2. Fundamentos de la Teoría del Valor Extremo.	19
2.1 Distribución Generalizada de Valor Extremo.....	20
2.2 Distribución Generalizada de Pareto.....	26
2.3 Función de Exceso Medio.....	30
2.4 Estimación de la prioridad u	32
2.5 Contraste de la bondad de ajuste del modelo.....	34
2.6 Método de Gumbel.....	36
2.7 Periodo de Retorno.....	37
3 Aplicación de la Teoría del Valor Extremo.	40
3.1 Descripción de la muestra.....	41
3.2 Ajuste de la Distribución Generalizada de Pareto.....	45
3.2.1 Elección del umbral óptimo.....	46
3.2.2 Estimación de parámetros.....	50
3.2.3 Verificación de la Bondad del Modelo.....	50
3.2.4 Inferencia y Predicción.....	51
3.3 Frecuencia de valores extremos por el Método de Gumbel.....	57

Conclusiones.

Bibliografía.

Anexos.

**AJUSTE DE UNA DISTRIBUCIÓN PARETO GENERALIZADA PARA EL
ANÁLISIS DE SINIESTRALIDAD**

Matthieu Martínez Huerta

Capítulo 1

Antecedentes.

Dada la creciente preocupación por la solvencia de las entidades aseguradoras, incluso a niveles internacionales, el objeto de esta investigación consiste en poner en manifiesto la utilidad de la teoría del valor extremo en la ciencia actuarial a través del estudio de uno de sus componentes: los siniestros extremos.

Desde un punto de vista práctico, esta teoría facilita la gestión del riesgo en las entidades aseguradoras al permitir realizar inferencia estadística sobre valores extremos en una población o en un proceso aleatorio y, por tanto, estimar la probabilidad de que ocurran sucesos más extremos que los recogidos hasta el momento.

La relevancia de la investigación reside en la posibilidad de inferir sobre la ocurrencia de eventos extremos en base a la información histórica disponible. La incertidumbre ante la ocurrencia de grandes siniestros se reduce si se consigue obtener una distribución límite que modele la siniestralidad más elevada y que permita conocer su probabilidad y periodos de retorno.

Con el fin de contrastar la utilidad y posibilidades que presenta la Teoría del Valor Extremo, se realizará un análisis aplicado a las bases de datos de siniestralidad de una entidad aseguradora, representativa en el ramo automotriz, cuya identidad se mantendrá anónima por razones de confidencialidad. La entidad que ha colaborado en esta investigación ha cedido una cartera de siniestros de 10 años de historia.

Para este ejercicio se asumirá como costo de siniestralidad tanto los pagos realizados en los siniestros ya terminados, como las provisiones o reservas de los siniestros pendientes por finalizar.

Una vez estimados los parámetros de las distribuciones que ajustan a los datos y comprobada la bondad del ajuste del modelo paramétrico estimado, el objetivo es la predicción e inferencia:

- i. Estimar la probabilidad de que un siniestro supere una determinada cuantía, especialmente si ésta es elevada.
- ii. Estimar el cuantil esperado dada una determinada probabilidad.

- iii. Inferir sobre el periodo de retorno o tiempo medio de espera hasta la ocurrencia de un evento extremo, y al mismo tiempo, el nivel de siniestralidad esperado para un determinado periodo de retorno.

Con carácter previo, se deben distinguir los conceptos evento extremo y valor extremo: un evento extremo proviene de una situación anómala (evento es el accidente o suceso dañoso) pero no tiene por qué dar lugar a un valor extremo, ya que el valor extremo es la cuantificación económica consecuencia del evento. Todos los valores extremos se han producido por la ocurrencia de un evento extremo.

Se considera como valor extremo al dato más grande o más pequeño en un conjunto de observaciones o de una variable aleatoria. El objeto de la Teoría del Valor Extremo es desarrollar métodos para cuantificar eventos extremos y sus consecuencias de un modo estadísticamente óptimo.

Desde la perspectiva de la presente investigación los valores extremos son los grandes siniestros y respecto a ellos existe un problema básico que es la insuficiencia de datos y la escasa experiencia para estudiar el comportamiento de los mismos.

Responsabilidad civil, reaseguro, incendio y catástrofes son algunas de las áreas susceptibles de generar grandes siniestros y, por tanto, su siniestralidad extrema es susceptible de ser estudiada a través de esta teoría.

1.1 Gestión de Riesgo.

Embrechts, Resnick y Samorodnitsky proporcionan una serie de conceptos respecto a la gestión del riesgo de crédito en una entidad financiera, que resultan válidos en la gestión de la siniestralidad del asegurador:

- Pérdida esperada (*Expected loss*): Pérdida que se asume como consecuencia de llevar a cabo una transacción en particular.
- Pérdida inesperada (*Unexpected loss*): Pérdida inusual pero predecible que la entidad debe ser capaz de absorber con el curso normal de su actividad.
- Pérdida extrema (*Stress loss*): Pérdida surgida de escenarios extremos posibles (aunque improbables) a los que las entidades deben sobrevivir, sin que se ponga en peligro la estabilidad y la solvencia de las entidades aseguradoras.

Conceptos similares a éstos se ponen de manifiesto en los documentos de trabajo relativos a Solvencia II, que establecen tres componentes del riesgo a los que la compañía debe hacer frente: Volatilidad, Incertidumbre y Eventos extremos. Estos tres componentes son la base de toda investigación actuarial interesada, por un lado, en estimar la prima pura como valor medio o esperanza matemática de la suma total de siniestralidad esperada de una cartera de riesgos y, por

otro lado, en estimar los recargos necesarios de seguridad a la prima de riesgo en función de la volatilidad y la incertidumbre del riesgo asegurado.

Estos dos primeros aspectos, volatilidad e incertidumbre, han sido ampliamente investigados a través de la teoría clásica del riesgo.

Sin embargo, el último de los riesgos, eventos extremos, tienen una historia bastante reciente en relación al resto. Actuarialmente comienza a considerarse la existencia de potenciales siniestros de baja frecuencia pero de graves consecuencias monetarias.

Su investigación es fundamental ya que para cualquier riesgo, sea o no de naturaleza actuarial, la ocurrencia de uno o más eventos extremos puede dar lugar a fluctuaciones mucho mayores de las que cabría esperar que aparecieran por fluctuaciones normales (tanto por volatilidad como por incertidumbre).

Particularmente para el seguro y concretamente para las entidades aseguradoras, el interés por la Teoría de los Valores Extremos es inherente al riesgo potencial de siniestros de elevada cuantía.

El objetivo de investigar el comportamiento de los valores extremos conlleva a asegurar la estabilidad y la solvencia de las entidades aseguradoras pese a la ocurrencia de pérdidas extremas, a través de un tratamiento técnico que ayude a la toma de decisiones.

Parte de los procedimientos técnicos para la gestión del riesgo, consisten en conocer y aplicar modelos estadísticos que midan los riesgos y permitan establecer óptimamente el volumen de capital necesario, el nivel adecuado de provisiones, la correcta tarificación, o las posibilidades de cesión del riesgo, etc.

Tal es la importancia sobre los valores extremos en el análisis de la solvencia de los aseguradores, que el proyecto de Solvencia II considera relevante su investigación y pone de manifiesto permanentemente su preocupación por la estabilidad de las entidades aseguradoras en relación a la siniestralidad más grave.

Para la medición de las situaciones extremas propone algunas medidas técnicas:

- Valor en riesgo (VaR – *Value at Risk*): Cuantil de la función de distribución que deja un determinado percentil $(1 - \alpha)\%$ a su izquierda y el resto a la derecha $(\alpha)\%$
- La desviación estándar de una distribución es una medida del grado de incertidumbre.
- Valor en Riesgo de cola (TVaR – *Tail Value at Risk*): Cuantil del VaR más la media que excede dicho cuantil.

La ventaja de usar la medida del TVaR como una medida del riesgo, con el objeto de valorar la solvencia, es proporcionar la suma total de pérdidas por encima de un cierto nivel de confianza $(1 - \alpha)\%$

De la misma manera, Solvencia II plantean varias aproximaciones para calcular la exposición al riesgo y factores específicos. Una de estas aproximaciones parte del ajuste de una distribución de probabilidad a las carteras de riesgo que determine la suma de capitales requeridos para soportar las responsabilidades futuras que se deriven de las mismas. De la distribución obtenida se toma una probabilidad elevada y se calcula el cuantil asociado a ese percentil, estableciendo como óptimo un nivel de confianza del 99,5% (ello implica un periodo de retorno de 200 años)

Bajo la Teoría del Valor Extremo se determina el modelo que ajuste la cola de la distribución, ya que para un mismo percentil, el valor del cuantil es más alto cuanto mayor sea el grueso de la cola de la distribución.

Ni todas las carteras de riesgo se rigen bajo un mismo modelo, (no todas presentan un comportamiento normal), ni un mismo modelo de riesgo es válido para el ajuste de toda una distribución.

En definitiva, la Teoría del Valor Extremo facilita, en primer lugar, el ajuste de un modelo a la cola de la distribución del riesgo y en segundo, el cálculo correcto de las medidas para optimizar la toma de decisiones.

Las distribuciones de cola gruesa muestran una mayor masa de probabilidad en la cola en comparación con otras distribuciones como la normal. Esto significa que los acontecimientos extremos que provienen de distribuciones de cola gruesa, ya sean éstos máximos o mínimos, ocurren con mayor frecuencia que bajo una hipótesis de normalidad.

Los riesgos sujetos a pérdidas infrecuentes pero cuantificables siguen distribuciones de cola gruesa que a su vez se corresponden, en su mayoría con distribuciones asimétricas.

Partiendo de una hipótesis de normalidad en la distribución de un fenómeno aleatorio, las estimaciones de los valores extremos se minimizarían. Por ello es necesario acudir a investigar el comportamiento propio de los extremos, separándolos de los valores centrales para ser estudiados por sí mismos y conseguir una mejor explicación de la realidad, sin asumir que todos los acontecimientos provienen de un comportamiento normal.

En términos actuariales, para determinar una prima que cubra los siniestros de cuantía muy elevada, donde en base a los datos históricos existen sólo unas pocas observaciones de cierta cuantía, ajustando un modelo paramétrico al conjunto de todos los siniestros (extremos y no extremos), la distribución no realizaría un buen ajuste a la cola de la distribución donde las observaciones son muy dispersas, de hecho, la optimización del ajuste implicaría, en base a la estadística clásica, la eliminación de los picos que contaminan los valores centrales.

Existen entonces dos posibilidades para las compañías aseguradoras: cubrir los grandes riesgos a través de sus propios fondos mediante la correcta estimación de la distribución de la cola o no asegurar elevadas cuantías debido a la escasa experiencia y ceder dichos riesgos a figuras como el reaseguro.

1.2 Reaseguro y Solvencia

El reaseguro es un acuerdo en virtud del cual el asegurador traslada a otro asegurador parte de los riesgos que originariamente había tomado a su cargo. Se trata de un seguro donde el primer asegurador toma la posición de asegurado frente a un segundo asegurador que se denomina reasegurador. Su finalidad es atenuar las obligaciones asumidas por el asegurador originario si ocurriesen los siniestros, repartiéndolas entre varios aseguradores y reduciendo así los tres componentes del riesgo mencionados con anterioridad: volatilidad, incertidumbre y eventos extremos.

El reaseguro ayuda a los aseguradores a salvaguardar su estabilidad financiera, por ejemplo, en un reaseguro no proporcional de exceso de pérdida, desde el punto de vista de la compañía cedente, la distribución de siniestros está truncada a la derecha (el asegurador tiene acotadas sus pérdidas).

Está claro que el reaseguro es una figura clave para mantener la estabilidad y la solvencia de las entidades aseguradoras, pero la transferencia del riesgo o el reparto de las consecuencias de un siniestro de elevada cuantía no evitan la ocurrencia del evento extremo.

Las entidades aseguradoras deben valorar las probabilidades de ocurrencia de extremos, su cuantía máxima probable, etc., con el fin de determinar si resulta conveniente su cesión al reaseguro o por el contrario la entidad podría ser capaz de asumirlos con capacidad propia.

En términos generales, solvencia es la capacidad de hacer frente a las deudas u obligaciones de pago, sin embargo, se puede hacer distinción entre solvencia estática y dinámica.

Se entiende por *solvencia estática* a la capacidad del asegurador para hacer frente a los compromisos adquiridos, esto es, la capacidad técnico – financiera para hacer frente a las obligaciones ya contraídas por la empresa.

El elemento básico para la cobertura de la solvencia estática de las entidades son las provisiones técnicas (excepto la provisión de estabilización), aquellos recursos que están destinados a atender el importe de las obligaciones futuras estimadas pero que aún no son exigibles, denominados fondos no libres.

La *solvencia dinámica* corresponde a los riesgos derivados de las fluctuaciones aleatorias de siniestralidad, los derivados de la gestión y derivados del mercado donde la empresa realiza su actividad.

La cobertura de la solvencia dinámica se realiza a través del margen de solvencia y el fondo de garantía (además de la propia provisión de estabilización), esto es, el nivel de fondos libres de los que dispone el asegurador. Estos fondos libres no son necesarios para atender obligaciones pendientes ni para cubrir el valor estimado de los compromisos futuros por razón de las pólizas suscritas hasta la fecha, sino que con esos fondos libres el asegurador debe ser capaz de atender desviaciones futuras desfavorables. El fondo libre del periodo anterior más las primas devengadas y los rendimientos de inversiones financieras, restando el costo y gasto de los siniestros, generará el fondo libre del periodo actual.

Mientras que la solvencia estática se identifica con la deuda ya contraída, la solvencia dinámica se asocia al riesgo subyacente en función del carácter intrínsecamente aleatorio de la actividad aseguradora expuesta a fluctuaciones y posibilidad de desestabilización de la entidad.

En caso contrario, la insolvencia vendrá cuando los fondos libres alcancen un nivel que se considere mínimo. Para mantener un adecuado grado de solvencia será condición necesaria que los ingresos, primas, no sean inferiores a los gastos, siniestralidad y gestión. A favor de la solvencia pueden añadirse recargos de seguridad a las primas, disminuir los gastos de gestión o mediante una mejor estimación de los costos medios de siniestralidad y, por tanto, ajustando de mejor manera las primas puras.

Además de la condición anterior, también será necesario establecer un *margen de solvencia* inicial a fin de minimizar el riesgo de insolvencia. Normativamente es requisito para las entidades aseguradoras establecer un margen de solvencia más allá de las provisiones técnicas para hacer frente a los compromisos contraídos.

1.3 Proceso de reservas.

Desde el punto de vista tradicional de la Teoría del Riesgo, en el análisis de la solvencia de las entidades aseguradoras existe sólo un tipo de riesgo: las fluctuaciones en la siniestralidad acumulada. Estas posibles fluctuaciones pueden afectar a la estabilidad del asegurador y por tanto definen su probabilidad de ruina.

Una entidad aseguradora será estable cuando la probabilidad de ruina sea suficientemente pequeña, y para ello, se debe considerar el efecto del reaseguro, del recargo de seguridad y demás elementos de solvencia que intervienen en el proceso asegurador.

Filip Lundberg introdujo los fundamentos de la teoría del riesgo en el campo de los seguros no vida, con un modelo capaz de describir las dinámicas básicas de una cartera de seguros no vida. Este modelo asume un proceso de Poisson para describir el número de siniestros y presenta un análisis en tiempo continuo.

Sea R_0 el volumen de reservas inicial o margen de solvencia en el periodo inicial. La cuantía de siniestralidad agregada para un determinado periodo será representada por una variable estocástica S y donde las primas cobradas estarán representadas por P' .

Las primas se suelen considerar como una variable determinista que viene definida por la prima pura de riesgo P y por el recargo de seguridad, normalmente expresado como un porcentaje λ de la prima pura de riesgo, siendo $P' = P \cdot (1 + \lambda)$ y el volumen de primas puras de riesgo coincide con la esperanza de la siniestralidad $P = E(S)$.

La entidad aseguradora será solvente si al final de un único periodo la siniestralidad no ha superado el volumen de reservas iniciales más las primas cobradas:

$$R_0 + P' > S$$

El volumen de reservas iniciales es conocido y el volumen de primas a cobrar se estima en base a la siniestralidad. Para conocer cuál debe ser el volumen inicial de reservas ante la variabilidad en los pagos, se plantea la estimación de R_0 tal que la probabilidad de que la siniestralidad sea superior a la suma de las primas y las reservas iniciales, sea inferior a un valor determinado $\varepsilon > 0$, el cual determina la probabilidad de insolvencia máxima aceptable:

$$P(S > R_0 + P') \leq \varepsilon$$

En la práctica, el corto plazo tiende a periodos mayores de tiempo. Aunque se trate de seguros con duración inferior a un año, siempre existe el componente de renovación y de continuidad en la actividad aseguradora. Una generalización del corto plazo lleva a una prolongación en el tiempo.

Para un proceso de riesgo que se desarrolla en un horizonte temporal $(0, t)$, sea $N_{(t)}$ el número de siniestros que pueden ocurrir hasta el momento $t \geq 0$, la probabilidad de que hasta el momento t hayan ocurrido k siniestros se puede expresar como:

$$p_{k,t} = P(N_{(t)} = k)$$

El modelo clásico de riesgo, cuando se analiza en tiempo continuo, viene definido por un proceso estocástico $\{R(t), t \geq 0\}$, donde:

$$R(t) = R_0 + P'_t - S(t) \quad t \geq 0,$$

Siendo $R(t)$ el nivel de reservas en el momento t , calculadas como la cuantía de las reservas en el momento 0 más la cuantía de las primas recibidas menos la cuantía pagada por los siniestros hasta el momento t .

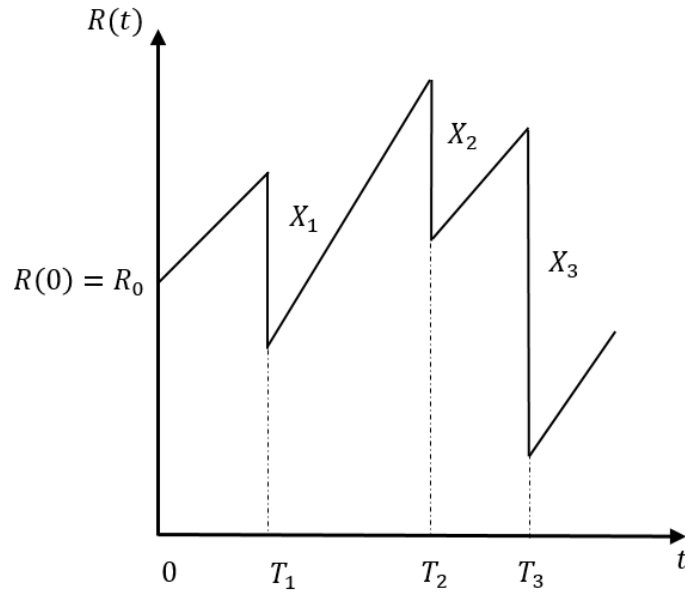
Así, $R_0 = R(0) \geq 0$ es el nivel inicial de las reservas, P' el monto de la prima, siendo por tanto P'_t el total de primas cobradas hasta el momento t , y $S(t)$ la siniestralidad agregada, es decir, el total de los siniestros ocurridos en $(0, t]$

$$S(t) = \sum_{i=1}^{N(t)} X_i$$

Siendo $N(t)$ el número de siniestros ocurridos hasta t , y X_i la cuantía del i -ésimo siniestro.

En el siguiente gráfico se representa una posible trayectoria del proceso. Las variables aleatorias $T_1, T_2, \dots$ son los momentos en los cuales sucede un siniestro. La pendiente del proceso es c en el caso en que no hay siniestros, y cuando $t = T_i$, $i = 1, \dots, n$ las reservas caen escalonadamente por la cuantía del siniestro i -ésimo, X_i .

Trayectoria de un proceso estocástico de reservas



En el modelo no se incluyen elementos como la inflación, rendimientos de la inversión, reparto de dividendos o gastos de carácter administrativo y gestión.

En el modelo clásico, el proceso de ocurrencia de siniestros $\{N(t), t \geq 0\}$ es un proceso de Poisson con media λ , donde $N(t)$ es el número de siniestros ocurridos en $(0, t]$. Siendo entonces $S(t)$ un proceso de Poisson compuesto, cumpliendo que $S(t) = 0$ si $N(t) = 0$

En el análisis del proceso de reservas, la insolvencia se produce cuando $R(t) < 0$.

La *probabilidad de insolvencia* es la probabilidad de que las reservas en un momento, entre 0 y t , alcancen un valor negativo condicionado a un nivel de reservas inicial R_0 , el cual comprende el margen de solvencia:

$$\psi(R_0, t) = P\{R(T) < 0, 0 < T \leq t \mid R(0) = R_0\} = P(T < t \mid R(0) = R_0)$$

Siendo $\psi(R_0, t)$ la función de distribución de la variable T .

La probabilidad de ruina ($0 \leq \psi(R_0) \leq 1$) es una función decreciente respecto a R_0 y tiende a cero cuando las reservas iniciales tienden a infinito.

No obstante, el problema para una entidad se produce si la reserva llegase por debajo de la barrera de ruina B_R , que sería aquella para la cual la situación de insolvencia es irreversible.

Para el periodo $(0, t]$ se define la probabilidad de ruina como la probabilidad de que las reservas alcancen un valor por debajo de la barrera o frontera de ruina:

$$\psi(R_0, t) = P\{R(t) < B_R \forall t > 0 \mid R(0) = R_0\}$$

Y sea la probabilidad de supervivencia el complemento de la probabilidad de ruina:

$$\varphi(R_0, t) = 1 - \psi(R_0, t)$$

Sin embargo, el problema importante es determinar un rango de variación para $R(t)$. Para ello se establece un límite superior R^S y un límite inferior R^I dentro del cual se encontrará la reserva final con una determinada probabilidad.

La probabilidad α de que $R(t)$ se encuentre al final del periodo $(0, t]$ por debajo del límite inferior, condicionado a que las reservas iniciales son R_0 es:

$$\alpha = P\{R(t) < R^I \mid R(t) = R_0\}$$

Se denomina capital en riesgo a la diferencia $R(t) - R^I$ que representa el capital que puede perderse como máximo para un nivel de confianza $(1 - \alpha)$

Sin embargo, el proceso de riesgo basado en una distribución única para toda la siniestralidad puede ser deficiente para estimar el volumen total de reservas o inferir la verdadera probabilidad de ruina, lo que lleva a la necesidad de emplear modelos eficientes que ajusten separadamente la cola del ajuste de la masa muestral (Anexo I).

1.4 Reaseguro de exceso de pérdida.

La teoría clásica del riesgo, basada en el teorema central del límite no proporciona modelos adecuados a la siniestralidad más elevada de las entidades aseguradoras.

Asumir que toda la siniestralidad se comporta normalmente o bajo un mismo modelo, es una restricción innecesaria si se puede aplicar la teoría del Valor Extremo. Del mismo modo, tampoco es válido asumir que una única distribución ajusta óptimamente todo el conjunto de siniestros.

Estimando la distribución del costo de siniestros y la distribución de frecuencia de siniestralidad, la teoría clásica del riesgo indicaría el importe de reservas necesario para hacer frente a sus obligaciones, dada una determinada probabilidad de ruina y un horizonte temporal.

El reaseguro de exceso de pérdida (*excess loss*), es aquel que cubre los siniestros que exceden una determinada prioridad P , fijada por la compañía cedente.

La entidad reaseguradora se hace cargo de la cuantía R_i de cada siniestro X_i si:

$$\begin{cases} X_i < P \rightarrow R_i = 0 \\ X_i > P \rightarrow R_i = X_i - P \end{cases}$$

La entidad cedente se hace cargo de la cuantía C_i de cada siniestro X_i si:

$$\begin{cases} X_i < P \rightarrow C_i = X_i \\ X_i > P \rightarrow C_i = P \end{cases}$$

Reasegurador

En base a su propia cartera de riesgo, el reasegurador debe conocer con exactitud qué tipo de siniestralidad grave asume y cuál es su modelo óptimo, ya que ello determinará el nivel de primas de riesgo adecuado a los riesgos aceptados.

La prima de tarifa, por parte del costo de la siniestralidad de una entidad cedente, se obtiene estimando la prima pura de riesgo a través de la cuantía de siniestralidad media esperada y un recargo de seguridad medido por la dispersión de la variable.

$$\text{Prima de tarifa} = E(R) + \lambda \cdot \sigma(R) \quad \lambda \in \mathbb{R}$$

Donde $E(R)$ es la prima pura determinada mediante la siniestralidad esperada que asumirá el reasegurador. A su vez, se determina mediante el producto del número de siniestros esperado y la esperanza de costo medio de cada uno de ellos, asumiendo independencia entre la frecuencia y el costo de la siniestralidad. El recargo de seguridad está formado a través de la desviación $\sigma(R)$ y un número real λ que mide la propensión al riesgo del reasegurador.

La prima pura de riesgo para el siguiente periodo puede ser estimada de un modo no paramétrico en función de la siniestralidad de periodos precedentes.

- **Estimación no paramétrica**

Sea N el número de siniestros que ocurrirán durante el periodo a estimar y X la cuantía de siniestralidad. La suma esperada de siniestralidad total S_N que ocurrirá en un periodo dado, será la esperanza del número de siniestros por la esperanza de la siniestralidad media de cada uno de ellos:

$$E(S_N) = E(N) \cdot E(X)$$

En base a esta expresión se puede determinar la prima neta mediante una estimación no paramétrica a partir del siguiente estimador insesgado:

$$E(S_N) \approx \frac{S_{N(T)}}{T}$$

Donde $S_{N(T)}$ es la suma total de siniestralidad ocurrida en T periodos.

- **Estimación paramétrica**

Si lo que se propone es estimar la prima pura de modo paramétrico, entonces:

$$PR_{(P)} = E(n) \cdot \int_P^{\infty} x \cdot dF(x) = E(N) \cdot \int_0^{\infty} (x - P) \cdot dF(x)$$

La prima del reaseguro de exceso de pérdida se expresará como el producto de la esperanza del número de siniestros y la esperanza de la cuantía de siniestralidad en exceso de la

prioridad P , en base a la Distribución Generalizada de Pareto. Donde $E(n)$ es el número esperado de siniestros por encima de P asumiendo una distribución de Poisson.

De este costo medio estimado el reasegurador cubrirá únicamente el exceso esperado sobre la prioridad, es decir, $[E(X | X > P) - P]$

Asumiendo que los momentos de ocurrencia y el tamaño de los siniestros satisfacen las condiciones de un proceso de Poisson(λ, W), entonces, λ es el número medio de siniestros en un periodo y W es la función de distribución de excesos por encima de la prioridad P (función Generalizada de Pareto).

La prima pura de riesgo para el siguiente periodo vendría dada por:

$$PR_{(P)} = \lambda \cdot m(W_{\xi, \beta, P}) = E(n) \cdot \int_P^{\infty} (x - P) \cdot w(x) = E(n) \cdot [E(X | X > P) - P]$$

Donde $m(W_{\xi, \beta, P})$ representa la esperanza de la función de distribución de Pareto Generalizada, con parámetros ξ, β y prioridad P .

Cedente

Desde el punto de vista del cedente, se presta interés en el riesgo de siniestralidad por encima de una prioridad a cambio de una cesión de primas o para mantener una capacidad financiera suficiente que permita atender la siniestralidad por encima de nivel de pérdida. En caso de cesión al reaseguro, se necesitará determinar el umbral de cesión y retención óptimos, es decir, el importe de la prioridad y precio del reaseguro.

Mayor cesión al reaseguro implica menor nivel de fondos libres necesarios (menor remuneración por costo de capital propio), pero mayor costo por la cobertura de riesgos graves (menores prioridades implican mayor cesión de primas). Y por el contrario, menor cesión al reaseguro (altas prioridades) implica mayor nivel de fondos libres necesario (mayor remuneración por costo de capital propio) pero menor costo por la cobertura de riesgos graves.

La entidad aseguradora debe valorar en términos de costo y tomar decisiones óptimas para conservar la futura estabilidad. Con el proyecto de Solvencia II, los requerimientos de capitales (*SCR: Solvency Capital Requirement*), tendrán una base más cuantificable apoyada en distribuciones de costo y percentiles de riesgo como el *VaR* y *TVaR*.

La teoría del Valor Extremo presenta los nuevos escenarios para la estabilidad del mercado y reaseguro en entornos económicos abordados por la investigación actuarial.

1.5 Limitaciones del Valor Extremo.

A pesar de la importancia de esta teoría, es necesario hacer notar que existen ciertas limitaciones que se deben tener en cuenta:

- Los modelos se desarrollan empleando leyes y teorías asintóticas, por lo tanto, los resultados no pueden ser tomados como exactos cuando se aplican sobre muestras finitas.
- Los modelos se desarrollan bajo circunstancias ideales. Sin embargo, la realidad bajo estudio no cumple siempre todas las hipótesis teóricas.
- El modelo restringido a los valores extremos puede suponer pérdida de información de aquellos valores elevados que se encuentren en el mismo bloque y que sean incluso superiores a los máximos de otros bloques.
- El modelo de los excesos que superan un umbral también genera un problema de elección entre varianza y sesgo. Umbrales bajos suponen modelos con mayor número de observaciones y por tanto minimizan la varianza, aunque pueden incrementar el sesgo por tomar como muestra valores no extremos. Umbrales elevados reducen el sesgo pero generan modelos con mayor varianza.

Capítulo 2.

Fundamentos de la Teoría del Valor Extremo.

El riesgo indica la siniestralidad de un evento y el asegurador tiene como objeto la cobertura económica de las consecuencias desfavorables derivadas de esos riesgos, midiendo la probabilidad de ocurrencia de los mismos.

Desde el punto de vista del seguro y de la previsión, el riesgo puede clasificarse en Riesgos de Masa y Riesgos de Cola. Mientras que los primeros son riesgos con consecuencias económicas bajas pero alta probabilidad de ocurrencia, los segundos son aquellos riesgos con baja frecuencia pero elevadas consecuencias económicas.

Una vez que el asegurador ha aproximado la distribución de la siniestralidad acumulada, puede determinar el volumen de primas necesario para cubrir los riesgos asegurados y el margen de solvencia que minimiza su probabilidad de ruina.

Los grandes riesgos se consideran valores extremos, y por serlo, resulta insuficiente la aplicación de la estadística clásica. De este modo surge la necesidad de recurrir a una metodología de cola de la distribución de probabilidad de la siniestralidad (*Heavy Tails* o *Long Tails*).

Si bien la distribución del daño total modela los riesgos de masa, permitiendo la aplicación de la Ley de los Grandes Números y el Teorema de Limite Central, la estimación de la siniestralidad acumulada debe completarse con otros estudios.

Las principales carencias de la teoría del riesgo se derivadas de los grandes siniestros; la carencia de datos, el tiempo hasta la declaración y el tiempo hasta la liquidación.

Cuando se intenta ajustar una distribución a las cuantías de siniestralidad surge el problema en los valores más altos, donde apenas existen datos. Puede haber un buen ajuste en los valores más bajos, pero no en la cola. La presencia o ausencia de un único gran siniestro tiene impactos sobre el comportamiento total de la distribución y puede desvirtuar la apreciación de la siniestralidad en un segmento.

El hecho de que haya ocurrido un siniestro y aún no se haya declarado afecta a la distribución del daño total. En determinados ramos, los siniestros de gran cuantía presentan un tiempo medio de declaración superior a la media.

En determinadas clases de riesgo, como la responsabilidad civil, aparece una complicación adicional en los grandes riesgos, el tiempo que transcurre desde que se produce y se declara un siniestro hasta que se liquida completamente. En siniestros de menor cuantía la liquidación se produce en el corto plazo, lo que facilita su estudio estadístico, mientras que en los grandes siniestros pueden transcurrir incluso años, lo que distorsiona la inferencia ante cambios en las condiciones económicas (inflación, tipos de interés, tipos de cambio etc.)

La investigación actuarial explica la necesidad de emplear otras metodologías estadísticas dirigidas a incrementar la información sobre los grandes riesgos. Estas nuevas técnicas se agrupan en la Teoría del Valor Extremo.

2.1. Distribución Generalizada de Valor Extremo.

La Teoría del Valor Extremo analiza la condición bajo la cual los extremos de una muestra aleatoria $X_1, \dots, X_n$, convergen hacia una distribución límite no degenerada cuando el tamaño de la muestra n tiende hacia infinito.

Para ello, se asume que las variables aleatorias X_i son independientes e idénticamente distribuidas respecto a una distribución dada F . Donde las variables aleatorias son estrictamente positivas y, por tanto, las funciones de distribución están truncadas en 0 por la izquierda.

Sean entonces $X_{1,n} \geq X_{2,n} \geq \dots \geq X_{k,n} \geq \dots \geq X_{n,n}$ un conjunto de variables aleatorias independientes e idénticamente distribuidas con una función de distribución común F [$F(x) = P(X_{i,n} \leq x)$], donde M_n representa el máximo de todas ellas, $M_n = (X_{i,n}) = \max(X_{1,n}, X_{2,n}, \dots, X_{k,n}, \dots, X_{n,n})$, se pretende conocer el comportamiento o distribución de los máximos.

Para obtener la distribución límite no degenerada es necesario reducir el máximo mediante la aplicación de una transformación lineal con coeficientes que dependan del tamaño de la muestra de modo análogo a una tipificación, es decir, seleccionando los apropiados parámetros $b_n > 0$ y $a_n \in \mathbb{R}$ se pretende llegar a la función de distribución del máximo.

El parámetro a_n es una medida de la tendencia central de la distribución de los valores extremos pero no es la media de la distribución. Del mismo modo, el parámetro b_n es una medida de la dispersión pero no la desviación estándar.

$$\frac{M_n - a_n}{b_n} \rightarrow X \text{ cuando } n \rightarrow \infty \quad \text{y} \quad P\left(\frac{M_n - a_n}{b_n} \leq x\right) \rightarrow G(x) \text{ cuando } n \rightarrow \infty$$

La Teoría del Valor Extremo centra su interés en la fluctuación del máximo muestral, en lugar de centrar la atención en las sumas acumuladas

$$\left(S(t) = \sum_{k=1}^n X_{k,n}\right)$$

La probabilidad, medida por la función de distribución del máximo, es:

$$P(M_n \leq x) = P(X_{1,n} \leq x; \dots; X_{n,n} \leq x) = \prod_{i=1}^n P(X_{i,n} \leq x) = F^n(x)$$

El comportamiento asintótico del máximo tiene que estar relacionado con la cola derecha de la distribución F , puesto que es allí donde ocurren los extremos.

Esta función de distribución del máximo viene dada por una serie de probabilidades conjuntas (ya que si el máximo es menor que x , es porque todos los demás valores ordenados también son menores que x), es decir, la función de distribución vendrá dada por el producto de probabilidades y, por tanto, por la convolución de la distribución F .

Esto es:

$$P\left(\frac{M_n - a_n}{b_n} \leq x\right) = F^n(a_n + b_n \cdot x)$$

Por el teorema de Fisher y Tippet sobre leyes límites de máximos, si existen las constantes $b_n > 0$ y $a_n \in \mathbb{R}$ y una función no degenerada G tal que:

$$\frac{M_n - a_n}{b_n} \xrightarrow{d} G \text{ cuando } n \rightarrow \infty$$

Entonces

$$G(x) = \lim_{n \rightarrow \infty} P\left(\frac{M_n - a_n}{b_n} \leq x\right) \quad x \in \mathbb{R}$$

$G(x)$ es una distribución asintótica derivada a partir de la función de distribución exponencial y representa la función de distribución estándar del valor extremo.

Ésta aproximación es la base de la Teoría clásica del Valor Extremo, que da respuesta a la primera cuestión anteriormente planteada y es que los máximos normalizados convergen hacia la Distribución del Valor Extremo Generalizada (*GEVD*).

La segunda cuestión viene referida a la forma de esa función de distribución límite $G(x)$ y a cómo se caracteriza el Máximo Dominio de Atracción (MDA) de G , es decir, cuáles son las funciones de distribución F que pertenecen al $\text{MDA}(G)$.

Para responder a ambas cuestiones primero se deben describir todas las distribuciones G que pueden surgir como límite de la expresión anterior y segundo se deben encontrar y caracterizar todas las distribuciones para las cuales a_n y b_n existen, de modo que la expresión anterior se sostenga.

Se dice que una distribución es *max-estable* cuando para cada $n = 2, 3, \dots$, existen los parámetros $b_n > 0$ y $a_n \in \mathbb{R}$, tal que: $F^n(a_n + b_n \cdot x) = F(x)$

Ya que F_n es la función de distribución de $M_n = \max(X_{1,n}, X_{2,n}, \dots, X_{k,n}, \dots, X_{n,n})$, donde X_i son variables independientes e idénticamente distribuidas con distribución común F , la propiedad de *max-estabilidad* la satisfacen aquellas distribuciones para las cuales la operación de tomar máximos muestrales da lugar a una distribución idéntica.

Como conclusión, las distribuciones *max-estable* son las únicas leyes límite para los máximos normalizados.

Las clases de distribuciones *max-estable* coinciden con las clases de todas las posibles leyes límite (no degeneradas) para los máximos (adecuadamente tipificados) de variables aleatorias independientes e idénticamente distribuidas.

Esto implica que una distribución sea *max-estable* si, y sólo si, es una distribución generalizada del valor extremo.

Del desarrollo del límite para la distribución asintótica, se concluye que la función de probabilidades acumuladas de los datos $F(x)$ se encuentra en el máximo dominio de atracción (MDA) de G , donde G puede tomar la forma de tres distribuciones:

- Tipo I: Distribución de Gumbel para colas medias.
- Tipo II: Distribución de Frechet para colas gruesas.
- Tipo III: Distribución de Weibull para colas cortas o suaves.

En cuanto a MDA, se dice que una variable aleatoria X pertenece al máximo dominio de atracción de una distribución del valor extremo si existen los parámetros $b_n > 0$ y $a_n \in \mathbb{R}$ de modo que $\frac{M_n - a_n}{b_n} \xrightarrow{d} G$ cuando $n \rightarrow \infty$ se cumpla.

Es decir, verifica que:

$$G(x) = \lim_{n \rightarrow \infty} P(M_n \leq b_n \cdot X + a_n) = \lim_{n \rightarrow \infty} F^n(b_n \cdot X + a_n) = G(X) \text{ con } X \in \mathbb{R}$$

Y se dice que $X \in MDA(G)$

En la siguiente tabla se resume la forma de las distribuciones límite para máximos extraídos de muestras dada una distribución inicial $F(x)$, (Anexo II).

<i>Distribución inicial $F(x)$</i>	<i>Distribución límite para los máximos $G(x)$</i>
Exponencial	<i>Tipo I (Gumbel)</i>
Gamma	<i>Tipo I (Gumbel)</i>
Normal	<i>Tipo I (Gumbel)</i>
Log – Normal	<i>Tipo I (Gumbel)</i>
Pareto	<i>Tipo II (Frechet)</i>
Cauchy	<i>Tipo II (Frechet)</i>
Burr	<i>Tipo II (Frechet)</i>
Log – Gamma	<i>Tipo II (Frechet)</i>
Uniforme	<i>Tipo III (Weibull)</i>
Beta	<i>Tipo III (Weibull)</i>

Las tres expresiones tienen en cuenta las variables estandarizadas, esto es, parámetro de localización nulo ($a_n = \mu = 0$) y parámetro de escala igual a la unidad ($b_n = \sigma = 1$), ya que:

$$G_{i,\alpha,a_n,b_n}(x) = G_{i,\alpha}\left(\frac{x - a_n}{b_n}\right)$$

Únicamente, las distribuciones de Frechet y Weibull están definidas en términos de un parámetro (parámetro cierto de forma) llamado índice de cola que indica el grado de convergencia a cero de la densidad de probabilidad, es decir, el grueso de la cola.

Cuanto mayor es el parámetro más se suaviza la cola, más rápidamente converge la función a la asíntota, mientras que si el parámetro se acerca a cero, mayor es el grueso de la cola de la distribución, más lentamente converge la función a cero.

Para poder realizar inferencia sobre la población objeto de estudio, deberán aplicarse técnicas que permitan la estimación de dicho parámetro.

A. Funciones de probabilidad acumulada y funciones de densidad

Tipo I: La distribución de Gumbel, ilimitada por los dos extremos, es una función doble exponencial cuya función de distribución y de densidad son, respectivamente, las siguientes:

$$\Lambda(x) = G_0(x) = \exp\{-\exp(-x)\}, \quad x \in \mathbb{R}$$

$$g_{0,\alpha}(x) = G_0(x) \cdot e^{-x}$$

Bajo esta distribución la cola decrece de forma exponencial.

El momento de primer orden es, para el caso de la distribución tipificada:

$$E(X) = \int_0^{\infty} (-\ln(x)) \cdot e^{-x} dx = \lambda$$

Introduciendo los parámetros de localización y escala:

$$E(X) = \bar{X} = a + \lambda \cdot b$$

Donde la varianza será:

$$\text{Var}(X) = \sigma^2(X) = \frac{\pi^2}{6}$$

E introduciendo los parámetros de localización y escala para la desviación:

$$\text{Var}(X) = \sigma(X) = \frac{b \cdot \pi}{\sqrt{6}}$$

Tipo II: Distribución de Frechet, acotada por la izquierda, sólo toma en cuenta los valores positivos. Utilizada en aplicaciones al seguro, reaseguro y finanzas.

$$\Phi(x) = G_1(x) = \begin{cases} 0, & x \leq 0 \\ \exp\{-x^{-\alpha}\}, & x > 0, \alpha > 0 \end{cases}$$

$$g_{1,\alpha}(x) = \alpha \cdot G_1(x) \cdot x^{-(\alpha+1)}$$

Bajo esta distribución la cola decrece de manera potencial (*power law*).

Tipo III: Distribución de Weibull. Acotada por la derecha, para valores positivos.

$$\Psi(x) = G_2(x) = \begin{cases} 1, & x > 0 \\ \exp\{-(-x)^{-\alpha}\}, & x \leq 0, \alpha < 0 \end{cases}$$

$$g_{2,\alpha}(x) = |\alpha| \cdot G_2(x) \cdot (-x)^{-(\alpha+1)}$$

En este último tipo, la cola permanece finita.

Las distribuciones de Frechet y Weibull tienden a la forma de la distribución de Gumbel cuando el parámetro de forma α tiende a ∞ y $-\infty$, respectivamente.

B. Funciones de cuantiles

Se denomina cuantil- p de una función de distribución G , al valor X_p tal que $G(X_p) = p$, es decir, al valor X_p que deja a su izquierda el percentil p de probabilidad.

Si una función de distribución G es continua y por tanto, estrictamente creciente, entonces la función cuantil es la inversa de la función de distribución G y se suele denotar como G^{-1} o bien como $Q(p)$.

Las tres distribuciones del valor extremo enunciadas se pueden generalizar en un modelo único a través de la reparametrización $\xi = \frac{1}{\alpha}$ y la elección de los apropiados parámetros de localización ($a_n = \mu$) y de escala ($b_n = \sigma$).

La representación única (*con el parametro ξ*) de las tres distribuciones tipificadas del valor extremo, da lugar a la expresión que conforma la Distribución del Valor Extremo Generalizada:

$$G_{\xi}(x) = \exp\left\{-(1 + \xi \cdot x)^{\frac{-1}{\xi}}\right\}, \quad (1 + \xi \cdot x) > 0$$

Donde la expresión toma diferentes formas en función del parámetro ξ :

Distribución de Gumbel si $\xi = 0$

$$G_{\xi=0}(x) = \exp\{-\exp(-x)\}, \quad x \in \mathbb{R}$$

Distribución de Frechet si $\xi = \frac{1}{\alpha} > 0$

$$G_{\xi>0}(x) = \exp\left\{-(1 + \xi \cdot x)^{\frac{-1}{\xi}}\right\} = \exp\{-(x')^{-\alpha}\}, \quad x > \frac{-1}{\xi}$$

Distribución de Weibull si $\xi = \frac{1}{\alpha} < 0$

$$G_{\xi<0}(x) = \exp\left\{-(1 + \xi \cdot x)^{\frac{-1}{\xi}}\right\} = \exp\{-(-x')^{-\alpha}\}, \quad x < \frac{1}{|\xi|}$$

La función de Distribución de Valor Extremo Generalizada incluyendo de forma explícita los parámetros de localización μ y de escala σ es:

$$G_{\xi;\mu,\sigma}(x) = \exp\left\{-(1 + \xi \cdot \frac{x - \mu}{\sigma})^{\frac{-1}{\xi}}\right\}, \quad \xi, \mu \in \mathbb{R} \quad \sigma > 0$$

De nuevo, la distribución toma las tres formas de las distribuciones básicas en función de cómo sean los parámetro característicos de la función.

Es decir, para $\xi > 0$ aparece ante la distribución de Frechet, que se encuentra entre las distribuciones de cola gruesa, para $\xi < 0$ será una distribución de Weibull, y para $\xi = 0$ la distribución corresponde a Gumbel.

La correspondiente función de densidad de la expresión generalizada es:

$$g_{i,\xi}(x) = G_i(x) \cdot (1 + \xi \cdot x)^{-(1+\frac{1}{\xi})}, \quad \begin{cases} i = 1 & \text{si } \xi > 0 & \text{Frechet} \\ i = 1 & \text{si } \xi < 0 & \text{Weibull} \end{cases}$$

$$g_{0,\xi}(x) = G_0(x) \cdot e^{-x}, \quad x \in \mathbb{R} \quad \xi = 0$$

Donde la función inversa o función de cuantiles es:

$$G_\xi^{-1}(p) = \frac{(-\ln(p))^{-\xi} - 1}{\xi}, \quad \xi \neq 0$$

En cuanto a sus momentos, están definidos en términos de la función integral Gamma:

$$E_\xi(X) = \frac{\Gamma(1 - \xi) - 1}{\xi}, \quad \xi < 1$$

Con varianza:

$$Var(X) = \frac{\Gamma(1 - 2 \cdot \xi) - \Gamma^2(1 - \xi)}{\xi^2}, \quad \xi < \frac{1}{2}$$

2.2 Distribución Generalizada de Pareto.

Sean $X_{1,n}, X_{2,n}, \dots, X_{n,n}$ un conjunto de variables aleatorias independientes e idénticamente distribuida, el denominado método P.O.T. (*peaks over threshold method*) permite estudiar cómo se distribuyen las observaciones una vez que éstas han superado una determinada cuantía, llamada prioridad o umbral y denotada por u , teniendo una relación directa con el reaseguro de exceso de pérdida o *excess of loss*, donde, dada una prioridad u el reasegurador debe pagar el exceso sobre la prioridad siempre y cuando el siniestro sea superior a u , es decir, el reasegurador paga $X - u$, condicionado a que $X > u$.

El método P.O.T. supone que los excesos, por encima de la prioridad, constituyen una secuencia de variables aleatorias independientes e idénticamente distribuidas que ocurren según los momentos de un proceso de Poisson y que las cuantías que exceden la prioridad y el momento del tiempo en el que ocurren son independientes.

Bajo este método, la principal dificultad se encuentra en la elección del umbral u apropiado, ya que la teoría clásica no aborda dicho aspecto.

La selección de un umbral óptimo está sujeta al problema de elección entre varianza y sesgo, es decir, si se incrementa el número de observaciones para formar las series de máximos disminuyendo el umbral, la estimación del índice de cola será más precisa (con menor varianza) pero sesgada, ya que algunas observaciones del centro de la distribución, no de la cola, se introducirán en la

serie. Por otro lado, si se reduce el número de observaciones con un umbral más elevado, se reduce el sesgo, pero la estimación del índice será más volátil al realizarse con un menor número de observaciones.

La función de distribución condicionada de los excesos por encima del valor u se denomina *función de distribución de exceso en u* [F^u] y viene dada por la siguiente expresión:

$$F^u = P(Y \leq y | X > u) = P(X - u \leq y | X > u); \quad \text{para } 0 \leq y \leq x_0 - u$$

Donde u representa el umbral, X representa la variable aleatoria de la cuantía total, Y la variable aleatoria del exceso sobre u , y x_0 el punto máximo de la distribución, finito o infinito, de modo que:

$$\begin{aligned} X &= Y + u \\ Y &= X - u \end{aligned}$$

Cuando el nivel de la prioridad y el número de valores que superan la prioridad son suficientemente grandes, la función condicionada ajusta tanto los valores que exceden u ($X_1, X_1, \dots, X_k$), como los excesos sobre u ($Y_1 = X_1 - u, Y_2 = X_2 - u, \dots, Y_k = X_k - u$), cambiando únicamente los parámetros de localización y escala.

Por tanto, la función de distribución condicionada proporciona la probabilidad de que la variable aleatoria de exceso Y sea menor a un determinado valor y condicionado a que el siniestro supere un determinado umbral u , y del mismo modo la probabilidad de que la variable aleatoria de la cuantía total del siniestro X sea inferior a un determinado valor x , condicionado a que el siniestro supere el umbral fijo u .

El punto máximo positivo de la función de distribución es definido como x_0 y el extremo inferior de la distribución o punto inicial es cero si se ajustan las cuantías totales y u si se ajustan los valores excedentes o excesos.

La probabilidad condicionada vendrá dada por:

$$\begin{aligned} F^u(x) = F^u(y + u) &= P(X \leq y + u | X > u) = P(X - u \leq y | X > u) = \frac{F(x) - F(u)}{1 - F(u)} \\ &\text{para } 0 \leq y \leq x_0 - u \end{aligned}$$

Y la función de supervivencia de exceso en u :

$$\begin{aligned} 1 - F^u(x) &= P(X > y + u | X > u) = P(X - u > y | X > u) = \frac{1 - F(x)}{1 - F(u)} \\ &\text{para } 0 \leq y \leq x_0 - u \end{aligned}$$

Un conjunto de resultados equivalentes a los desarrollados para la Distribución del Valor Extremo Generalizada, permite describir el comportamiento de estas extremas observaciones cuando exceden un umbral.

El teorema de Pickands-Balkema-de Haan (Anexo III) proporciona fundamentos teóricos y demuestra que, cuando la prioridad u toma un valor suficientemente grande y bajo condiciones de máximos dominios de atracción, los datos por encima de ese punto siguen una distribución límite que es la función de Distribución Pareto Generalizada $W_\xi(x)$ (GPD por sus siglas en inglés), bajo previa elección del umbral óptimo y la estimación del parámetro ξ .

Por tanto, la función condicionada de probabilidad acumulada estandarizada cuando se ajustan los excesos sobre el umbral u , es la siguiente:

$$W_\xi(x) = 1 - \left(1 + \xi \cdot \frac{x}{\beta}\right)^{-\left(\frac{1}{\xi}\right)} \quad \text{para} \quad \begin{cases} x > 0 & \text{si } \xi > 0 \\ 0 < x < \frac{1}{|\xi|} & \text{si } \xi < 0 \end{cases}$$

Donde $\beta = \sigma + \xi(u - \mu)$

Esta función indica la probabilidad de que una observación sea menor a x , una vez que se ha superado el umbral u , donde ξ es el parámetro característico de la distribución o el índice de cola (cuanto mayor sea el parámetro más gruesa es la cola).

La función condicionada de supervivencia estandarizada (complemento de la función condicionada de distribución) mide la probabilidad de que la variable aleatoria supere un determinado valor condicionada a ser superior al umbral fijo:

$$P(X - u > x \mid X > u) = 1 - W_\xi(x) = \left(1 + \xi \cdot \frac{x}{\beta}\right)^{-\left(\frac{1}{\xi}\right)}$$

El teorema implica que si los máximos se aproximan a la distribución Generalizada del Valor Extremo (G), entonces, los excesos por encima de un umbral se aproximan a la Distribución Generalizada de Pareto (W), obteniendo la siguiente relación:

$$W_\xi(x) = 1 - \log[G_\xi(x)]$$

Derivando la función de distribución se obtiene la función de densidad estandarizada:

$$w_\xi(x) = (1 + \xi \cdot x)^{-\left(1 + \frac{1}{\xi}\right)} \quad \text{para} \quad \begin{cases} x \geq 0 & \text{si } \xi > 0 \\ 0 \leq x < \frac{1}{|\xi|} & \text{si } \xi < 0 \end{cases}$$

Por último, la función de cuantiles estandarizada de la distribución Generalizada de Pareto queda de la siguiente manera:

$$W_\xi^{-1}(p) = \frac{(1 - p)^{-\xi} - 1}{\xi} = X_p \quad \text{para} \quad \xi \neq 0$$

Submodelos de la Distribución Pareto Generalizada.

La distribución Generalizada de Pareto unifica tres submodelos en uno único a través de la reparametrización $\xi = \frac{1}{\alpha}$

Las tres distribuciones que pertenecen a su familia son la distribución exponencial cuando el parámetro ξ tiende a 0, la distribución de Pareto cuando el parámetro ξ es mayor que 0 y la distribución Beta cuando el parámetro ξ es negativo.

A continuación se muestran la función de distribución $W_\xi(x)$, función de densidad $w_\xi(x)$ y función de cuantiles $W_\xi^{-1}(p)$ para cada caso:

- Distribución Exponencial: $\xi = 0$

$$\begin{aligned} W_{\xi=0}(x) &= 1 - e^{-x} & x \geq 0 \\ w_{\xi=0}(x) &= e^{-x} & x \geq 0 \\ W_{\xi=0}^{-1}(p) &= -\ln(1-p) \end{aligned}$$

Para valores de x inferiores a cero, las funciones de distribución y de densidad Exponencial son iguales a cero.

- Distribución Pareto: $\xi > 0$ con la reparametrización $\xi = \frac{1}{\alpha}$, $\alpha > 0$

$$\begin{aligned} W_{\xi,\alpha}(x) &= 1 - x^{-\alpha} & x \geq 1 \\ w_{\xi,\alpha}(x) &= \alpha \cdot x^{-(1+\alpha)} & x \geq 1 \\ W_{\xi}^{-1}(p) &= (1-p)^{-\left(\frac{1}{\alpha}\right)} \end{aligned}$$

Para valores de x inferiores a uno, las funciones de distribución y de densidad Pareto son iguales a cero.

- Distribución Beta: $\xi < 0$ con la reparametrización $\xi = \frac{1}{\alpha}$, $\alpha < 0$

$$\begin{aligned} W_{\xi,\alpha}(x) &= 1 - (-x)^{-\alpha} & -1 \leq x \leq 0 \\ w_{\xi,\alpha}(x) &= |\alpha| \cdot (-x)^{-(1+\alpha)} & -1 \leq x \leq 0 \\ W_{\xi}^{-1}(p) &= -(1-p)^{-\left(\frac{1}{\alpha}\right)} \end{aligned}$$

Para valores de x inferiores a -1, las funciones de distribución y de densidad Beta son iguales a cero y para valores de x mayores a cero es igual a 1. Notar que, si α toma el valor -1, se tiene la distribución uniforme en el intervalo $[-1,0]$

Estimación paramétrica

La estimación de los parámetros a través de máxima verosimilitud cumple características óptimas y tienden a ser insesgados, sin embargo, la función de máxima verosimilitud se resuelve a través de un algoritmo iterativo que requiere de una estimación inicial para los parámetros.

Por otra parte, Smith, R.L. demostró que a pesar de la violación de las condiciones de regularidad estándar, las buenas propiedades de los estimadores de máxima verosimilitud sólo se sostienen cuando $\xi > -\frac{1}{2}$, pero no en el caso de $\xi \leq -\frac{1}{2}$.

Esta condición implica excluir la estimación del parámetro en las distribuciones de cola muy corta. No obstante, la mayoría de las distribuciones encontradas en seguros y finanzas no son limitadas por la derecha, obteniendo $\xi > 0$, y generando buenos estimadores mediante el procedimiento de máxima verosimilitud.

Para la Distribución Generalizada de Pareto y suponiendo que los valores $x_1, x_2, \dots, x_k$ son los k excesos sobre la prioridad u , para $\xi \neq 0$, su función logarítmica máximo verosímil es:

$$L(\beta, \xi) = -k \cdot \ln(\beta) - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^k \ln\left(1 + \frac{\xi \cdot x_i}{\beta}\right)$$

Donde $\left(1 + \frac{\xi \cdot x_i}{\beta}\right) > 0$, $i = 1, 2, \dots, k$

Para estimar el parámetro característico ξ de la Distribución Generalizada de Pareto, el estimador de Drees-Pickands toma las k observaciones más altas de la muestra

$$\hat{\xi}_{k,n} = \frac{\ln\left(\frac{X_k - X_{2k}}{X_{2k} - X_{4k}}\right)}{\ln(2)} \quad \xi \in \mathbb{R}$$

Donde X_i representa la i -ésima observación más alta y el parámetro k representa el número de datos empleados en el proceso de estimación, es decir, las k mayores observaciones en el estimador.

2.3 Función de Exceso Medio

La Distribución de Pareto Generalizada es una función de distribución condicionada a que los valores excedan un cierto umbral u .

La función de distribución de exceso sobre $u[F^u]$ definida por:

$$F^u(y) = P(Y \leq y \mid X > u) = P(X - u \leq y \mid X > u)$$
$$F^u(x) = P(X \leq x) = F^u(y + u) = P(X - u \leq y \mid X > u) = \frac{F(x) - F(u)}{1 - F(u)}$$

Para $0 \leq y \leq x_0 - u$

Ofrece una aproximación de la distribución para valores extremos jerarquizados por encima de un valor u .

La siguiente cuestión a determinar es el valor de exceso medio dado por la esperanza condicionada de $X-u$ dado $X > u$.

La cuantía $e(u)$, es el denominado valor medio en exceso del umbral u , que en un contexto asegurador puede ser interpretado como la siniestralidad esperada una vez que los siniestros superan la prioridad u , cuando la cuantía posible de los siniestros es ilimitada.

$$e_{(u)} = E(X - u | X > u) = \int_0^{\infty} x \cdot dF^u(x)$$

La función de distribución del exceso tiene una relación directa con el reaseguro de exceso de pérdida, ya que el reasegurador sólo cubre aquella cuantía que excede la prioridad seleccionada, es decir, a través de la transferencia de riesgo, el reasegurador cubre $X-u$ condicionado a que $X > u$.

A su vez, el valor medio del exceso, permitiría obtener la esperanza de siniestralidad por encima de la prioridad u , es decir, la prima pura de riesgo del reaseguro de exceso de pérdida.

Dejando el umbral como variable independiente y el valor de exceso medio como variable dependiente se obtiene, $e_{F(u)}$, la función media de excesos (*mean excess loss function*), o bien función de exceso medio.

La función de exceso medio representa los valores de siniestralidad medios esperados en función del umbral elegido. Esta función es uno de los puntos básicos a la hora de verificar la hipótesis de Pareto Generalizada en la cola de una distribución y se presenta como una herramienta gráfica de utilidad para la discriminación en las colas, es decir, la primera aproximación al comportamiento de la cola se realiza a través de la observación de la forma de esa función de exceso medio.

Para datos provenientes de una distribución exponencial, la función de exceso medio es constante e independiente del valor u . El gráfico de exceso medio mostrará, aproximadamente, una línea horizontal.

Si provienen de una distribución con cola más débil que la exponencial (por ejemplo, la distribución de Weibull), la función será decreciente a medida que incremente la prioridad (tendencia descendente). Entonces, una función de exceso medio decreciente indica que la función presenta un punto final por la derecha y, por tanto, los grandes valores sólo juegan un papel minoritario.

Por el contrario, si los datos provienen de una distribución con cola más gruesa que la exponencial (por ejemplo, la distribución de Pareto Generalizada), la función de exceso medio será creciente a medida que incremente la prioridad (tendencia ascendente)

En general, la función de exceso medio para series de cola gruesa se localiza entre la función de exceso medio de una distribución exponencial y la función de exceso medio de la Distribución Generalizada de Pareto, que es lineal con pendiente positiva y tiende a infinito para valores grandes de u .

Si la distribución de procedencia de los datos se ajusta a una distribución Exponencial, la función de exceso medio será constante e independientemente del valor del umbral. Si la distribución de procedencia es de cola gruesa la función de exceso medio será lineal con pendiente positiva pero, si por el contrario, provienen de una distribución de cola suave la función de exceso medio será lineal con pendiente decreciente con tendencia a cero.

Las funciones de exceso medio para la distribución de Pareto generalizada y sus submodelos son:

- Distribución de Pareto Generalizada:

$$e_{w_\xi}(u) = \frac{1 + \xi \cdot u}{1 - \xi}, \quad \text{para } \begin{cases} u > 0 & \text{si } 0 \leq \xi < 1 \\ 0 < u < \frac{-1}{\xi} & \text{si } \xi < 0 \end{cases}$$

Esta es una función lineal de u con pendiente $\frac{\xi}{1-\xi}$. De aquí se desprende que el parámetro característico de la distribución puede ser aproximado a través de la pendiente de esta función.

- Distribución Exponencial (GP_0): $e_{w_0}(u) = 1, \quad u > 0$
- Distribución de Pareto ($GP_1, \alpha > 1$): $e_{w_1}(u) = \frac{u}{\alpha-1}, \quad u > 1$
- Distribución Beta ($GP_2, \alpha < 0$): $e_{w_2}(u) = \frac{u}{\alpha-1}, \quad -1 \leq u \leq 0$

2.4 Estimación de la prioridad u .

Dado que el gráfico de exceso medio para la distribución Generalizada de Pareto es lineal y tiende a infinito con pendiente positiva, el ajuste a través de la distribución Generalizada de Pareto es razonable si se toma como prioridad óptima el valor a partir del cual el gráfico de exceso medio toma una forma lineal con pendiente positiva.

Para ello, es necesaria la representación de la función empírica de exceso medio, obtenida a partir de la muestra disponible y que permitirá ir conociendo la distribución de los datos y su naturaleza.

Dada una muestra ordenada de forma descendente ($X_{1,n} \geq X_{2,n} \geq \dots \geq X_{k,n} \geq \dots \geq X_{n,n}$), la función empírica de exceso medio viene dada por la siguiente expresión:

$$\hat{e}_n(u) = \frac{\sum_{i=1}^n (X_i - u)^+}{\sum_{i=1}^n (X_i > u)}$$

En el numerador se encuentra la suma de los excesos sobre la prioridad y en el denominador el número de valores que cumplen la condición de ser superiores al umbral, determinándose así la media aritmética de los valores que exceden del umbral u .

Sin embargo, a efectos prácticos, resulta de interés que los propios datos de la muestra actúen como prioridad, esto es $u = X_{k+1}$. Si se toman como umbrales los propios datos, la función de exceso medio empírica que resulta será la media aritmética de los k mayores valores:

$$E_{k,n} = \hat{e}_n(X_k) = \frac{\sum_{i=1}^k X_i}{k}, \quad k = 1, 2, \dots, n-1$$

La representación de la función empírica frente a los umbrales da lugar al gráfico empírico de exceso medio, donde la función empírica $\hat{e}_n(u)$ es la variable dependiente y u la variable independiente. Las coordenadas de los puntos a representar son $(u, \hat{e}_n(u))$.

En el caso en que $u = X_{k+1}$, entonces, el gráfico de exceso medio se representará mediante k , es decir, con el par de valores $(k, E_{k,n})$ para $k = 1, \dots, n-1$, o bien mediante X_{k+1} con las coordenadas $(X_{k+1}, E_{k,n})$ para $k = 1, \dots, n-1$.

Hay que tener en cuenta, que bajo la primera opción, tomando como variable independiente los valores de k , a medida que se avanza en el eje de abscisas hacia la derecha, el valor empírico de exceso medio se determina con un mayor número de observaciones y, por tanto, con menor prioridad (a medida que aumenta k , disminuye la prioridad).

Mientras que bajo la segunda opción, se representa la función frente a las propias observaciones. En este caso, a medida que se avanza en el eje de abscisas, aumenta la prioridad para determinar el valor empírico de exceso medio. Es decir, ambos gráficos deben tener una apariencia prácticamente simétrica y poder establecer comparaciones con las funciones teóricas y distinguir así el modelo a seleccionar para el ajuste de los datos.

Intuitivamente, cuanto mayor es el valor del umbral u , más dispersa es la media aritmética de los valores que lo superan, ya que cada vez el número de observaciones es más reducido. Además, el cálculo de esta media puede estar influenciado por un número pequeño de datos atípicos que no sean realmente representativos de la tendencia general de los extremos.

Con la intención de fortalecer el procedimiento, Beirlant, J., Teugels, J.L. proponen las siguientes adaptaciones para los valores de $E_{k,n}$:

- Utilizar los valores de exceso medio restringidos con $p \in (0,1)$:

$$T_{k,n}^p = \frac{\sum_{i=[p_k]+1}^k X_i}{k - [p_k]} - X_{k+1}$$

Siendo $[p_k]$ el mayor entero no superior a k .

- Reemplazar la media de los datos que son mayores que X_{k+1} por el valor de la mediana generalizada, obteniendo así los valores de exceso de la mediana.

$$M_{k,n}^p = X_{[pk]+1} - X_{k+1}$$

Función cuadrática de Exceso Medio

La función cuadrática de exceso medio se define por:

$$s(u) = E[(X - u)^2 \mid X > u]$$

Esta función también juega un papel muy importante si nos referimos al reaseguro de exceso de pérdida con prioridad u , cuando se pretende fijar la prima en base al principio de varianza o desviación típica. Otras aplicaciones de $s(u)$ se emplean en la estimación de la cola cuando se usan estimadores de momentos.

$s(u)$ se define solamente cuando X cuenta con una varianza finita:

$$s(u) = 2 \cdot \frac{\int_u^{x^+} (a - u)(1 - F(a)) da}{1 - F(u)}$$

La función cuadrática de exceso medio empírica, se determina tomando $u = X_{k+1}$ de modo que $s(u) = S_{k,n}$:

$$S_{k,n} = \frac{1}{k} \sum_{i=1}^k (X_i - X_{k+1})^2 \quad k = 1, 2, \dots, n-1$$

El comportamiento de la función cuadrática de exceso cuando X tiende a infinito, proporciona información de la cola de la distribución a través de la comparación con la distribución exponencial, ya que su función cuadrática de exceso medio de nuevo es constante e independiente del valor de la prioridad.

2.5 Contraste de la bondad de ajuste del modelo.

El objeto de toda modelación es la inferencia para, a partir de eventos históricos, predecir futuras observaciones.

La inferencia y previsión de la siniestralidad más elevada permite a la entidad aseguradora tomar medidas de protección, tales como un incremento del margen de solvencia o de los fondos no comprometidos o bien buscar la cesión del riesgo mediante el reaseguro.

Para poder tomar la decisión adecuada el asegurador necesita estimar, predecir el nivel esperado de siniestros graves a través de un modelo preciso que ajuste el comportamiento histórico de la siniestralidad y permita inferir el futuro.

Un método no paramétrico que compara la distribución subyacente de la que provienen los datos con el modelo ajustado, son los gráficos de cuantil – cuantil y los gráficos de probabilidad – probabilidad.

Mediante el gráfico cuantil – cuantil se compara la función de cuantiles empírica con la función de cuantiles teórica (inversa de la función de distribución) de distintos modelos paramétricos. A través del mismo se puede saber si la distribución que fue tomada como referencia ajusta o no a la distribución real de los datos.

Si el gráfico de cuantil-cuantil tiene un comportamiento lineal, se puede afirmar que los datos muestrales se ajustan a la distribución tomada como referencia.

Se denomina cuantil- p de una función de distribución F , al valor X_p tal que $F(X_p) = p$, es decir, se denomina cuantil- p al valor X_p que deja a su izquierda el percentil p de probabilidad. El cuantil- p teórico viene dado por la función de cuantiles del modelo seleccionado (inversa de la función de distribución) estimando los percentiles a través de la expresión $p = \frac{r}{(n+1)}$ donde $r = 1, \dots, n$ tal que $F^{-1}(p) = Q(p) = X_p$

Dada una muestra ordenada de forma ascendente ($X_{1,n} < X_{2,n} < \dots < X_{r,n} < \dots < X_{n,n}$), el gráfico de cuantil-cuantil representa los valores muestrales, ordenados de menor a mayor, como variable dependiente y como variable independiente, el valor de los cuantiles teóricos, de modo que las coordenadas del gráfico que permiten comprobar la bondad del ajuste son:

$$(F_n^{-1}(p), X_{r,n})$$

Si el gráfico muestra un comportamiento lineal entonces $F_n^{-1}(p) \approx X_{r,n}$ y la distribución modelada presenta un buen ajuste muestral, de lo contrario, los datos presentan una cola más larga o más gruesa que dichas distribuciones y deberá buscarse otro ajuste.

La linealidad se puede contrastar rápidamente de forma visual, o de forma más precisa, por medio de coeficientes de correlación, mediante el cálculo del error cuadrático medio o mediante rectas de regresión.

Dadas dos variables aleatorias X y Y , cuando existe una función continua no decreciente que las relaciona $Y = g(X)$, se mantiene también la misma relación entre los correspondientes cuantiles de sus respectivas distribuciones, $Q_y(p) = g(Q_x(p))$.

Cuando se poseen dos muestras de datos ($X_{1,n} < X_{2,n} < \dots < X_{r,n} < \dots < X_{n,n}$) y ($Y_{1,n} < Y_{2,n} < \dots < Y_{r,n} < \dots < Y_{n,n}$) distribuidas como X y Y respectivamente, el gráfico de los datos ordenados de modo decreciente, con coordenadas $(X_{r,n}, Y_{r,n})$, proporciona una vista previa de la función g que hay que tomar para transformar la distribución X en la distribución Y . Si el gráfico de cuantiles empíricos se muestra lineal y cercano a la bisectriz, se puede decir que las distribuciones de X y Y no son diferentes, de lo contrario una de las dos distribuciones será más dispersa que otra.

Si en el gráfico de cuantil-cuantil se representaban los valores teóricos frente a los valores empíricos de la distribución en uso, en el gráfico de probabilidad-probabilidad, por el contrario, se representan las probabilidades:

$$(p_r, F_n(X_{r,n}))$$

En el eje de abscisas se presentan los percentiles, determinados tal y como se indicó anteriormente, y en el eje de ordenadas la probabilidad correspondiente a los valores muestrales bajo la distribución elegida como referencia.

Si la muestra se modela mediante la distribución de referencia, la línea que ajuste los puntos representados en un eje de coordenadas debe mostrarse lineal. Y al igual que en el gráfico de cuantil – cuantil, el coeficiente de correlación asociado al ajuste lineal de los datos en el gráfico será el indicador de la bondad del ajuste. Si los puntos representados muestran desviaciones respecto a la línea recta, se puede decir que la distribución no ajusta el conjunto de datos.

Mediante el gráfico de distribución empírica, la función de distribución muestral $\hat{F}_n(c)$ para series de datos univariados $[X_1, X_2, \dots, X_n]$ es el número relativo de valores X_i que son menores o iguales a un valor c de modo que:

$$\hat{F}_n(c) = \frac{1}{n} \cdot \sum_{i \leq n} 1(X_i \leq c)$$

La sumatoria indica el número de valores que cumplen la condición de ser inferiores o iguales al valor c , mide la proporción o frecuencia relativa de aquellos datos que no son mayores que un valor particular c .

Esta función consigue aproximar la función de distribución acumulada, es decir

$$F(c) = P(X \leq c)$$

Para contemplar esta aproximación de forma visual, se puede realizar la representación en un eje de coordenadas con los valores $(c, \hat{F}_n(c))$, para obtener el gráfico de la distribución empírica y proporcionar una primera aproximación de la distribución para el grueso de los datos, sin embargo, no puede emplearse para extrapolar más allá de la muestra.

2.6 Método de Gumbel

El recuento o frecuencia de las observaciones se modela mediante distribuciones discretas, de igual modo, el recuento de extremos se realiza en una escala nominal, de forma que las observaciones son un conjunto de eventos clasificados como extremos en grupos de observaciones o que exceden un criterio fijo. El objetivo será estimar el número medio de valores extremos que ocurrirán en los próximos periodos.

Partiendo de la observación de una serie de datos históricos se determina en cuántas de las próximas observaciones se excederá algunos registros anteriores.

Dado un conjunto ordenado de observaciones independientes e idénticamente distribuidas $X_{1,n} \geq X_{2,n} \geq \dots \geq X_{k,n} \geq \dots \geq X_{n,n}$, donde:

- n número de observaciones históricas o tamaño de la muestra de referencia
- k la posición u orden de una observación histórica en una jerarquía descendente
- N el tamaño de la muestra para la que se quiere pronosticar.
- h el número de observaciones que serán iguales o excederán el valor $k - \text{ésimo}$ en los próximos N ensayos o periodos de tiempo. Esto es, número de valores extremos esperado.

Sea $w(h)$ la distribución de h , donde k, n, N son parámetros que dan respuesta a la probabilidad p de que un nuevo dato sea menor que el $k - \text{ésimo}$ valor de los datos pasados, es decir, de que h casos de las N observaciones futuras igualen o excedan el valor $k - \text{ésimo}$ de las n observaciones pasadas.

Ante una situación binomial, se llega a la función de densidad de la frecuencia de los valores extremos:

$$w(h) = \frac{\binom{N+n-k-h}{n-k} \binom{h+k-1}{k-1}}{\binom{N+n}{n}} = \frac{k \binom{n}{k} \binom{N}{h}}{(N+n) \binom{N+n-1}{h+k-1}}$$

Donde

$$0 \leq h < N \quad y \quad 1 \leq k \leq n$$

Y cuya función de densidad es:

$$W(x) = \sum_{h=0}^x w(h)$$

Esta distribución de excedentes es análoga a la distribución binomial excepto en que la probabilidad del suceso es una cuantía variable.

2.7 Periodo de Retorno

Además del número de siniestros extremos y el volumen esperado de los mismos, se pretende conocer el momento estimado de ocurrencia entre los eventos analizados hasta el siguiente valor extremo, llamado periodo de retorno, tiempo de espera o intervalo de recurrencia.

La variable aleatoria es el momento de ocurrencia de un próximo excedente por encima de un valor específico Z en un horizonte indefinido de tiempo. Para estudiar el periodo de retorno, las

observaciones se hacen por intervalos regulares de tiempo y el experimento se interrumpe cuando el valor Z es excedido.

Sea una secuencia de variables aleatorias $X_1, X_2, \dots$ con una función de distribución común F , el momento r –ésimo en el que el valor X_t excede el valor Z se expresa como:

$$\tau_r = \min\{t > \tau_{r-1} : X_t > Z\} \quad r \geq 1$$

Por tanto, la probabilidad de que el excedente ocurra en el ensayo t y supere por primera el valor Z se puede expresar como una distribución geométrica con parámetro p .

Su función generatriz de momentos se define como:

$$G_t(m) = p \cdot (e^{-m} - (1 - p))^{-1}$$

De donde la media del primer momento en que se superará Z , también llamado periodo medio de retorno, es:

$$E(\tau_1, Z) = \frac{1}{1 - F(Z)} = \frac{1}{p} = T$$

Si un suceso tiene una probabilidad de ocurrencia p , entonces tiene que ocurrir $\frac{1}{p}$ ensayos para que el suceso ocurra una vez más.

Esta ecuación puede invertirse para extraer el valor Z en un periodo dado T , es decir, el valor Z tal que el número medio de excedentes por encima de Z en un periodo de tiempo de longitud T sea igual a 1. Por tanto, se entiende que para cada unidad de tiempo (año, mes,...) existe una observación:

$$E \left[\sum_{i \leq T} I(X_i \geq Z) \right] = 1$$

Así, se puede conocer el valor que ocurrirá con un tiempo medio de espera T a través de la siguiente expresión:

$$Z(T) = F^{-1} \left(1 - \frac{1}{T} \right)$$

Siendo una variable aleatoria geométrica con parámetro $p = \frac{1}{T}$ con varianza $V(\tau_1, Z) = T^2 - T$ y mediana $\mu(\tau_1, Z) = \frac{\ln(2)}{-\ln(1 - \frac{1}{T})} \approx \ln(2) \cdot T - \ln(\sqrt{2})$

El periodo de retorno toma mayor valor si las observaciones se realizan en intervalos de tiempo equidistantes, es decir, si suponemos que se realizan mediciones mensuales y que el máximo

mensual es el objeto de estudio. En este caso, el periodo de retorno es el número de meses que, en promedio, habría que esperar antes de que un siniestro de una magnitud específica ocurra otra vez.

Capítulo 3.

Aplicación de la Teoría del Valor Extremo.

Con el fin de contrastar la utilidad y posibilidades que presenta la Teoría del Valor Extremo, se realizará un estudio empírico aplicado a una bases de datos de siniestralidad de una entidad aseguradora en el ramo automotriz, cuya identidad se mantendrá oculta por razones de confidencialidad, denominándola “Entidad A”.

La relevancia de la investigación reside en la posibilidad de inferir sobre el acontecimiento de eventos extremos en base a la información histórica disponible. La incertidumbre ante la ocurrencia de grandes siniestros se reduce si se consigue obtener una distribución límite que modele la siniestralidad más elevada y que permita conocer su probabilidad y periodos de retorno.

Para este ejercicio se asume como costo de siniestralidad tanto los pagos realizados en los siniestros ya terminados, como las provisiones o reservas de los siniestros pendientes de finalización.

La influencia del paso tiempo y, concretamente el efecto de la inflación sobre la muestra, se ha tomado en cuenta en los siniestros ya terminados en la fecha de realización del análisis estadístico. En los siniestros pendientes de liquidación, la constante actualización de provisiones y el continuo flujo de pagos, llevan al desconocimiento de la inflación que afecta a cada siniestro y se asume que su actual costo está actualizado y expresado en unidades monetarias actuales.

Cuando se trata de estudiar sucesos con baja frecuencia y altos periodos de retorno, resulta importante tomar una base histórica suficiente pero sin incurrir en muestras que no se ajustan al presente. La entidad aseguradora que ha colaborado en esta investigación ha cedido una cartera de siniestros de 10 años de historia. Las bases de datos utilizadas permiten distinguir entre dos tipos de siniestros:

- El costo del siniestro está formado por todos los pagos realizados, actualizados con la inflación desde la fecha de liquidación total hasta la fecha del análisis. Ocurredos dentro del periodo de observación ya finalizados.
- El costo se genera por los pagos ya efectuados y/o las reservas de pagos futuros estimados. Este costo se considera actualizado con la inflación en el momento del estudio. Ocurredos dentro del periodo de observación pendientes de finalización.

Para cada siniestro se dispone de:

- El costo total de cada siniestro compuesto por la suma de los pagos realizados y/o las provisiones dotadas. De manera general, las cuantías serán denominadas en unidades monetarias (u.m.).
- Fecha de ocurrencia del siniestro.
- Fecha de declaración del siniestro.
- Fecha de finalización del siniestro (en caso de estar completamente liquidado).

Para realizar la equivalencia monetaria las cuantías nominales se han transformado a sus equivalentes con una tasa de inflación, representando la situación para n periodos, que en este caso corresponde con los meses desde la fecha de liquidación hasta diciembre de 2014.

3.1 Descripción de la muestra.

En la siguiente tabla se presentan, para cada uno de los 10 años objeto de estudio, el número de siniestros registrados, el costo total de siniestralidad en u.m. y el costo promedio individual en u.m. nominales:

Siniestralidad					
<i>Año</i>	<i>Número de siniestros</i>	<i>Costo total nominal u. m</i>	<i>Variación anual</i>	<i>Coste promedio nominal u. m</i>	<i>Variación anual</i>
<i>1</i>	21,721	15,719.58		0.72	
<i>2</i>	24,710	19,537.95	24.29%	0.79	9.26%
<i>3</i>	27,636	22,028.75	12.75%	0.80	0.81%
<i>4</i>	30,352	25,688.54	16.61%	0.85	6.18%
<i>5</i>	33,915	25,702.54	0.05%	0.76	-10.46%
<i>6</i>	35,322	27,177.50	5.74%	0.77	1.53%
<i>7</i>	39,326	30,774.35	13.23%	0.78	1.71%
<i>8</i>	41,258	33,503.91	8.87%	0.81	3.77%
<i>9</i>	41,713	36,334.96	8.45%	0.87	7.27%
<i>10</i>	42,175	39,187.78	7.85%	0.93	6.67%
Total:	338,128	275,655.86		0.82	

En cuanto a la siniestralidad respecto a las pólizas en cartera:

Pólizas en cartera

<i>Año</i>	<i>Número de siniestros</i>	<i>Variación anual</i>	<i>Número de pólizas</i>	<i>Variación anual</i>	<i>Relación siniestros/pólizas</i>
<i>1</i>	21,721		140,760		15.43%
<i>2</i>	24,710	13.76%	156,174	10.95%	15.82%
<i>3</i>	27,636	11.84%	173,166	10.88%	15.96%
<i>4</i>	30,352	9.83%	182,520	5.40%	16.63%
<i>5</i>	33,915	11.74%	190,344	4.29%	17.82%
<i>6</i>	35,322	4.15%	197,514	3.77%	17.88%
<i>7</i>	39,326	11.34%	209,322	5.98%	18.79%
<i>8</i>	41,258	4.91%	221,772	5.95%	18.60%
<i>9</i>	41,713	1.10%	236,892	6.82%	17.61%
<i>10</i>	42,175	1.11%	248,856	5.05%	16.95%
Total:	338,128		1,957,320		17.28%

Como consecuencia de la trayectoria estable y consolidada en el mercado, la entidad A posee una frecuencia de siniestralidad sobre sus pólizas acotada entre el 15% y el 19%. La media ponderada por el número de siniestros es de 17.28% sin existir fuerte desviación respecto a la media.

Esta información expresada en bruto, revela que un estudio agrupado de siniestralidad no permite extraer conclusiones particularizadas para cada grupo de riesgo o para cada tramo de costo. Un análisis detallado de siniestralidad demanda segmentación, que a efectos de esta investigación se requiere por tramo de costo, para identificar, aislar y modelar el comportamiento de los siniestros más graves.

A continuación se examina el costo del peor siniestro ocurrido cada mes durante los 10 años estudiados. Se busca detectar si existen comportamientos estacionales o si los meses de peor siniestralidad son los mismos año tras año.

Maximos mensuales

<i>Mes</i>	<i>Año 1</i>	<i>Año 2</i>	<i>Año 3</i>	<i>Año 4</i>	<i>Año 5</i>	<i>Año 6</i>	<i>Año 7</i>	<i>Año 8</i>	<i>Año 9</i>	<i>Año 10</i>
1	567.56	476.49	346.36	566.93	705.60	893.76	375.69	682.22	640.99	531.02
2	571.06	370.86	474.95	501.41	<u>2,720.20</u>	308.98	612.29	447.86	451.08	633.64
3	267.75	531.65	875.70	<u>1,197.70</u>	453.74	627.83	511.00	<u>841.89</u>	<u>729.47</u>	573.65
4	685.58	444.50	532.14	530.25	525.21	704.06	416.36	391.23	520.87	<u>848.54</u>
5	527.52	324.52	752.15	749.70	551.18	240.87	538.30	614.25	556.92	820.33
6	545.86	876.40	443.38	690.83	536.83	816.41	453.18	350.91	492.10	258.65
7	323.96	354.41	850.29	1,187.76	544.04	617.68	700.14	536.97	716.31	606.90
8	<u>680.26</u>	772.45	650.51	916.44	717.99	539.14	856.24	454.30	488.18	631.05
9	296.31	691.39	391.09	609.35	736.33	349.93	598.85	742.28	401.17	621.67
10	507.15	496.44	<u>1,979.39</u>	459.06	547.33	544.95	649.60	322.07	572.67	592.27
11	575.82	248.78	836.64	544.88	1,008.98	599.83	568.40	484.26	402.50	530.81
12	654.78	<u>1,070.23</u>	473.06	317.87	437.15	<u>1,371.79</u>	<u>883.96</u>	622.16	647.85	452.34
Max:	685.58	1,070.23	1,979.39	1,197.70	2,720.20	1,371.79	883.96	841.89	729.47	848.54

No parece observarse que el máximo anual de entre los máximos mensuales se concentre en los mismos meses año tras año.

Destaca el mes de marzo y diciembre, en 3 años su costo máximo ha estado entre el máximo anual.

Sin embargo, la limitación de esta primera aproximación es la pérdida de valores elevados de siniestralidad, no máximos mensuales, pero posiblemente superiores al máximo mensual de otro mes. Por ello se analizan los 10 y los 20 máximos anuales.

Tomado los 10 mayores siniestros anuales para los 10 años de objeto de estudio, nos indicarán la existencia de tendencias y/o estacionalidad:

- Los diez mayores siniestros anuales superan en todos los casos las 350 u.m. y sólo 6 siniestros de los 100 superan la cuantía de 1,050 u.m.
- El 94% restante de los siniestros se localiza entre las 350 y las 1,050 u.m. de la siguiente manera; el 63% entre 350 y 700 y el 21% entre 700 y 1,050.

El promedio y desviación estándar de los 10 máximos anuales se muestran en la siguiente tabla:

<i>Año</i>	<i>Promedio u. m.</i>	<i>Desviación estandar</i>
1	606.41	41.15
2	715.47	122.00
3	826.35	303.50
4	804.86	170.60
5	873.53	474.15
6	746.97	173.00
7	663.53	82.45
8	622.86	79.40
9	600.81	59.65
10	662.34	65.65
Total	712.31	204.90

En términos totales, el costo promedio de los 10 máximos anuales ha sido de 712 u.m. con una desviación estándar de 205 u.m., lo que supone un coeficiente de variación de 28.77%.

Este costo promedio resulta bastante estable, incluso decreciente a partir del quinto año, nos indican unos máximos anuales bastante centrados en torno al promedio.

A pesar de observarse ciertas diferencias respecto a las frecuencias de máximos en los distintos meses, se medirá mediante un contraste de hipótesis si dichas diferencias son estadísticamente significativas o no. El contraste empleado está basado en la distribución Chi-Cuadrada de una sola cola y llevará a determinar si existe o no una mayor concentración de siniestralidad en algunos meses, bimestres o trimestres.

El contraste de hipótesis formulado ha sido el siguiente:

Hipótesis nula (H_0): Todos los meses presentan la misma probabilidad.

Hipótesis alternativa (H_1): Algún mes presenta distintas probabilidades.

El estadístico experimental se ha construido del siguiente modo:

$$X^2 = \sum_{i=1}^n \frac{\theta_i - e_i}{e_i} \approx X_{n-1}^2$$

Donde θ_i son las frecuencias observadas y e_i las frecuencias esperadas, estimadas a través del cociente $\frac{1}{n}$ donde $n = 12$ para el contraste mensual, $n = 6$ para el bimestral y $n = 4$ para el trimestral.

Para comprobar el resultado del contraste de hipótesis, se mide la probabilidad del estadístico experimental obtenido a través de comparar las frecuencias observadas y las frecuencias teóricas. Si la probabilidad del estadístico siguiendo una distribución Chi-Cuadrada es uno, o cercana a la unidad, se acepta la hipótesis nula y se rechaza la hipótesis alternativa:

$$\begin{aligned}
X_{10 \max}^2 &= 0.0464 \approx X_{12-1}^2 \rightarrow p = 1 \\
X_{10 \max}^2 &= 0.0404 \approx X_{6-1}^2 \rightarrow p = 1 \\
X_{10 \max}^2 &= 0.0344 \approx X_{4-1}^2 \rightarrow p = 0.98
\end{aligned}$$

Las probabilidades del estadístico experimental son la unidad, lo que lleva a aceptar la hipótesis nula y asumir que ni las frecuencias mensuales ni las trimestrales son significativamente distintas en ninguna de las compañías.

Repitiendo el proceso con 20 máximos anuales en lugar de 10, se busca diferencias significativas entre una mayor concentración de siniestralidad en algunos meses, bimestres o trimestres a lo largo de los 10 años de estudio, obteniendo:

$$\begin{aligned}
X_{20 \max}^2 &= 0.0458 \approx X_{12-1}^2 \rightarrow p = 1 \\
X_{20 \max}^2 &= 0.0344 \approx X_{6-1}^2 \rightarrow p = 1 \\
X_{20 \max}^2 &= 0.0114 \approx X_{4-1}^2 \rightarrow p = 1
\end{aligned}$$

Por tanto, estos resultados indican que estadísticamente se debe aceptar la hipótesis nula y rechazar la hipótesis de que haya distinta probabilidad para algunos meses, bimestres o trimestres del año.

No obstante, no se puede negar que existe recurrencia en la concentración de siniestros graves en determinados periodos del año, la máxima siniestralidad se concentra en los mismos meses del año y en determinadas épocas. Los contrastes de hipótesis realizados únicamente contrastan si dichas diferencias son o no significativas, y a pesar de no serlo, no deja de ser un hecho la concentración y estacionalidad de la siniestralidad más grave.

El mes de agosto percibe la siniestralidad más grave, mientras que febrero se muestra como el mes con menor frecuencia de siniestros extremos.

Bimestralmente, siguen siendo los meses de verano (julio y agosto) los que mayor frecuencia de siniestros graves presentan. Se puede concluir que periodos vacacionales de verano son la principal época de ocurrencia de siniestralidad extrema. Sin ser de la misma forma marzo y abril o diciembre y enero.

En términos trimestrales es persistente la mayor frecuencia de grandes siniestros en el tercer trimestre, mientras que semestralmente es el segundo.

Esta regularidad demuestra que, aunque las diferencias entre meses, bimestres y trimestres no son estadísticamente significativas, sí existe un comportamiento estable.

3.2 Ajuste de la Distribución Generalizada de Pareto

Con el objeto de tomar decisiones financieras óptimas, las entidades aseguradoras y reaseguradoras necesitan inferir la siniestralidad extrema futura y el ajuste de la distribución de Pareto Generalizada facilita la inferencia de aquellos valores que exceden altas prioridades.

Desde el punto de vista de la entidad aseguradora, el interés se centra en conocer su capacidad financiera para cubrir la siniestralidad grave esperada, frente a la posibilidad de cesión al reaseguro, ya que, es el reaseguro de exceso de pérdida (*excess loss*) el que mantiene una relación directa con la distribución de los siniestros que exceden un determinado umbral.

Desde el punto de vista del reasegurador, el interés por la modelación de los extremos se centra en determinar la esperanza de siniestralidad por encima de una prioridad y la prima pura de dicho reaseguro.

El ajuste de la distribución de Pareto Generalizada a una muestra de observaciones implica dos cuestiones; la elección del umbral óptimo u y la estimación de sus parámetros.

Las hipótesis consideradas en el estudio son:

- Los valores extremos y su siniestralidad son variables aleatorias.
- Las observaciones son independientes e idénticamente distribuidas sin variar sustancialmente a lo largo del tiempo.
- El número de observaciones del cual se extraen los extremos es alto para sustentar las aproximaciones probabilísticas.

Teniendo como objetivos:

- i. Encontrar el umbral óptimo por encima del cual la distribución de Pareto Generalizada pueda ser ajustada a los excesos sobre dicho valor, utilizando la función empírica de exceso medio.
- ii. A través de estimadores que minimicen el Error Cuadrático Medio, se estimarán los parámetros del modelo ajustado a la muestra.
- iii. Comprobar la bondad del ajuste a la distribución apoyándose en métodos gráficos.
- iv. Realizar inferencia en base al modelo estimado. El modelo es una distribución condicionada y la inferencia realizada está condicionada a que un siniestro supere el umbral elegido, determinando las probabilidades marginales se consigue la distribución no condicionada.

3.2.1 Elección del umbral óptimo

En virtud del teorema de Pickands-Balkema-de Haan, la distribución generalizada de Pareto ajusta aquellos valores que exceden un determinado umbral cuando éste umbral u toma un valor suficientemente grande. La primera dificultad radica en determinar ese umbral a partir del cual es posible ajustar la distribución de Pareto Generalizada.

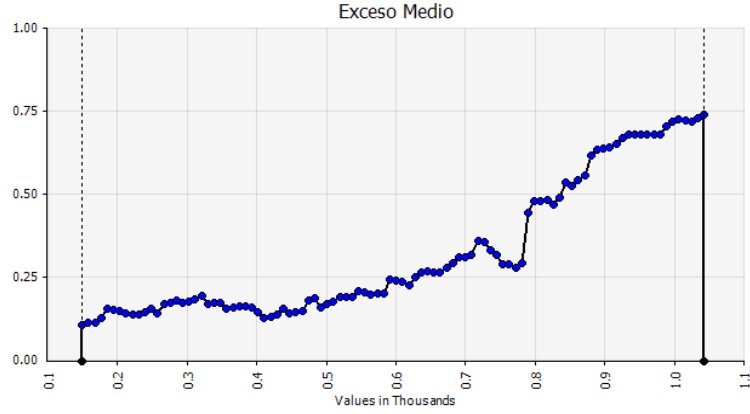
La primera aproximación ha consistido en la elección de una prioridad que dejara por encima alrededor de (10%) 1,000 observaciones que muestren el comportamiento de la función empírica de exceso medio y determinar el promedio de los valores que exceden la prioridad u .

Dada una muestra ordenada descendentemente $X_{1,n} \geq X_{2,n} \geq \dots \geq X_{k,n} \geq \dots \geq X_{n,n}$ la función empírica de exceso medio se determina a través del cociente entre la suma de los excesos sobre la prioridad y el número de valores que cumplen la condición de ser superiores a u . No obstante, en términos prácticos resulta de interés tomar como umbrales los propios datos de la muestra, esto es, $u = X_{k+1}$ donde la función de exceso medio empírica que resulta es la media de los k mayores valores muestrales.

$$E_{k,n} = \hat{e}_n(X_k) = \frac{\sum_{i=1}^k X_i}{k} - X_{k+1} \quad k = 1, 2, \dots, n-1$$

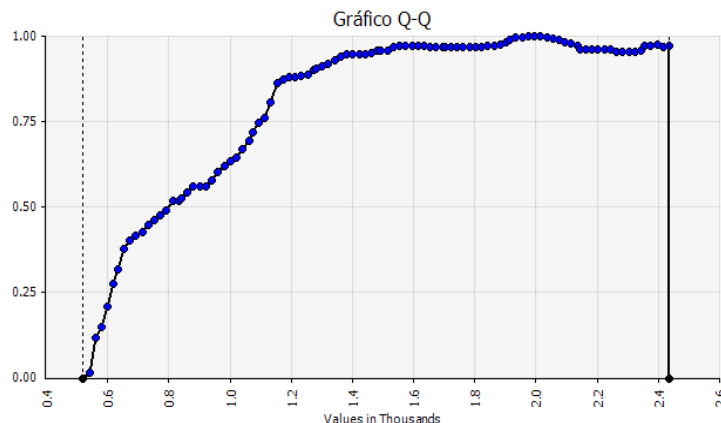
Para representar el gráfico de exceso medio se toma la función empírica $E_{k,n}$ como variable dependiente y los valores de la prioridad $u = X_{k+1}$ como variable independiente.

La representación de la función de exceso medio tomando un primer umbral de 175 u.m. (por tomar una muestra de 1,000 observaciones) es la siguiente:



En el eje de abscisas se representan las cuantías de los siniestros y en el eje de ordenadas el correspondiente valor de exceso medio donde se observa que entre los valores 140 y 490 la función de exceso medio es relativamente constante, mientras que aproximadamente en el valor 525 (525.37) la función empieza a ser creciente indicando que pueden provenir de una distribución de cola gruesa.

El siguiente paso es comprobar si los datos presentan una cola gruesa, como la distribución exponencial, o si por el contrario los valores extremos no son significativos y la distribución presenta un punto final por la derecha.



En el eje de abscisas se presentan los cuantiles empíricos o muestrales (basado en el gráfico empírico de exceso medio se toman aquellos que superan el umbral de 525 u.m.) frente a los cuantiles teóricos, representados en el eje de ordenadas y generados a través de la función inversa de la distribución exponencial.

Como en todo gráfico de cuantil-cuantil, la bondad del ajuste de la distribución elegida viene dada por una línea recta, mientras que la curvatura en la misma indica que la muestra presenta una cola más gruesa que la de la distribución exponencial.

Mediante la Razón de Verosimilitud (Likelihood ratio test) se contrasta el ajuste de la distribución exponencial a las observaciones, hipótesis nula, y por tanto $\xi = 0$ frente a la hipótesis alternativa $\xi \neq 0$.

La hipótesis de cola exponencial es rechazada si los valores resultantes son cercanos a 0, mientras que valores cercanos a 1 indicarían una aceptación de la hipótesis nula, es decir, el parámetro de la distribución es nulo y la cola decae exponencialmente.

Para la muestra de observaciones que superan 525 u.m. se obtiene:

$$P_{value} = 0.000253$$

Estas tres evidencias, de existencia de cola gruesa, indican la posibilidad de modelar la cola de las observaciones que superan 525 u.m. a través de la distribución generalizada de Pareto, sabiendo que el parámetro característico de la distribución será superior a cero.

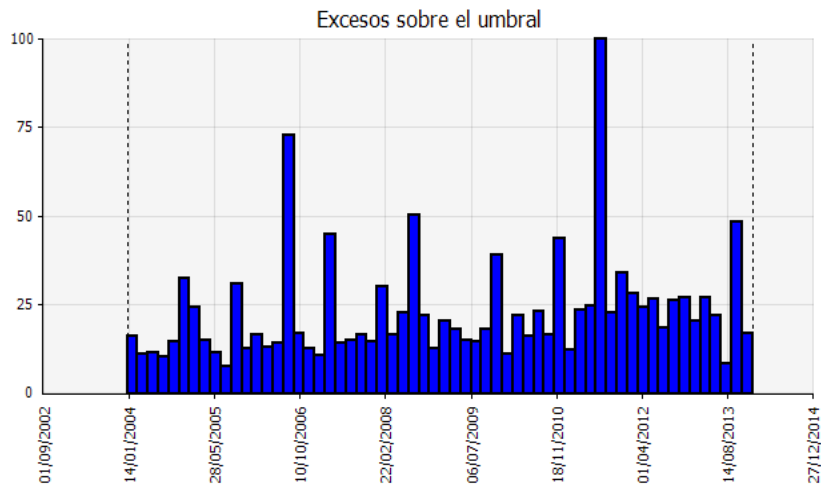
Por tanto, los estadísticos descriptivos de la muestra de siniestros que superan el umbral son:

<i>percentiles (q)</i>	<i>Valor para el percentil (X_q)</i>	<i>Número de valores por encima de $X_q(k)$</i>
0.25	556.92	$k = 651$
0.50	603.96	$k = 434$
0.75	704.13	$k = 217$
0.90	876.4	$k = 84$

Tamaño de la muestra = 875
Media = 680.47
Mínimo = 525.14
Máximo = 2,720.19

Varianza = 69,200.56
Desviación estándar = 263.06
Mediana = 603.96

Su representación gráfica frente al momento de ocurrencia junto con una regresión lineal a los datos indica una tendencia ligeramente descendente en el costo de los mismos:



El contraste de hipótesis formulado es el siguiente:

Hipótesis nula H_0 : Todos los meses presentan la misma probabilidad.

Hipótesis alternativa H_1 : Algún mes presenta distintas probabilidades.

El estadístico de prueba construido es:

$$X^2 = \sum_{i=1}^{12} \frac{\theta_i - e_i}{e_i} \approx X_{12-1}^2$$

Donde θ_i son las frecuencias observadas y e_i las frecuencias esperadas, estimadas a través del cociente $\frac{1}{12}$. El estadístico obtenido es una Chi-Cuadrada con 11 grados de libertad.

La probabilidad del estadístico experimental obtenido es igual a 1, por lo que se debe aceptar la hipótesis nula y se rechaza la hipótesis alternativa.

En cuanto a la posible estacionalidad, aunque se observa una menor masa de siniestros en los meses de febrero, abril y noviembre, en el resto de meses el reparto no muestra diferencias demasiado significativas.

3.2.2 Estimación de parámetros de la Distribución Pareto Generalizada

Hasta el momento es sabido que los siniestros más graves presentan cola gruesa y que la distribución de Pareto puede ajustar óptimamente a partir de 525 u.m. La muestra que supera esta prioridad, presenta una tendencia muy ligeramente descendente y una estacionalidad estadísticamente no significativa.

La siguiente fase es estimar los parámetros característicos de la distribución de ajuste y, tras diversos análisis para modelar la siniestralidad grave de la entidad A, han sido tres las decisiones tomadas para optimizar el ajuste:

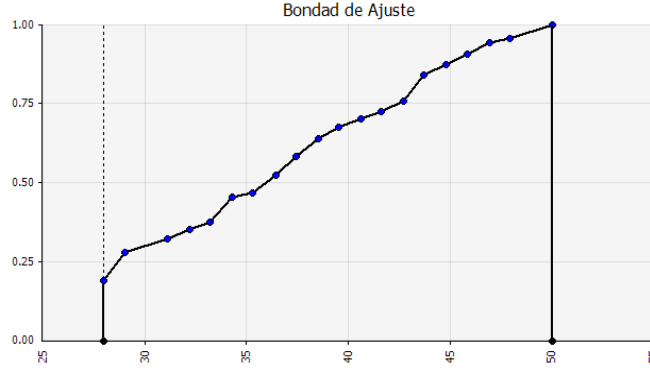
- De todos los posibles métodos de estimación de los parámetros, el estimador de Drees-Pickands minimiza el Error Cuadrático Medio en el ajuste de la distribución Generalizada de Pareto a la muestra disponible.
- Con el ajuste de la distribución al logaritmo natural de la muestra, en lugar de a las propias observaciones, se consigue mayor exactitud en el modelo.
- El número de extremos empleados para la estimación de los parámetros ha sido el número de valores excedentes $k=826$ que deben tomarse en la estimación para optimizar los parámetros del modelo, lo que corresponde con un umbral de 525.37 u.m.

Obteniendo así los valores $\xi = 0.3702$, $\beta = 0.2718$ y $\mu = 6.4952$

3.2.3 Verificación de la Bondad del Modelo

Para contrastar la bondad de ajuste se realiza el gráfico de cuantil-cuantil de la distribución obtenida una vez estimados los parámetros. En el eje de abscisas se presentan los cuantiles teóricos, es decir, el logaritmo natural de las observaciones y en el eje de ordenadas los cuantiles teóricos, $\hat{W}_{\xi, \beta, \mu}^{-1}(p) = x_p$.

El ajuste se considera óptimo cuanto más se aproxime el valor teórico a la observación muestral, cuanto más se aproxime a la bisectriz mayor será la bondad del ajuste realizado:



Se comprueba la ligera pérdida de la linealidad en el tramo de valores donde la función de exceso medio empírica deja de ser ascendente, no obstante, la mayor desviación del cuantil teórico respecto del cuantil empírico es de 0.0213 donde el coeficiente de determinación del ajuste lineal realizado (R^2) es el 97.94%

El Error Cuadrático Medio de dicho ajuste es de 0.0021 y como medida de contraste, el promedio de las desviaciones relativizadas (PDR) es de 0.24%, donde

$$PDR = \frac{\sum_{i=1}^n \left(\frac{|x_i^\theta - x_i^t|}{x_i^\theta} \right)}{n}$$

Y $|x_i^\theta - x_i^t|$ representa la diferencia en valores absolutos entre el cuantil teórico estimado con la distribución ajustada x_i^t y las observaciones muestrales x_i^θ

Por lo que se puede confiar en la bondad del ajuste de la distribución estimada a los extremos disponibles.

3.2.4 Inferencia y Predicción

Inferencia y predicción condicionadas:

A partir de la estimación de parámetros realizada y confiando en la bondad del ajuste, se puede realizar inferencia y predicción a partir de las funciones de distribución, de densidad y de cuantiles teóricos.

Debido a que el ajuste se ha realizado con el logaritmo natural de los datos, se necesitará de su transformación para obtener la predicción en u.m.

La función de probabilidades acumuladas, a partir de los parámetros estimados, se puede expresar como:

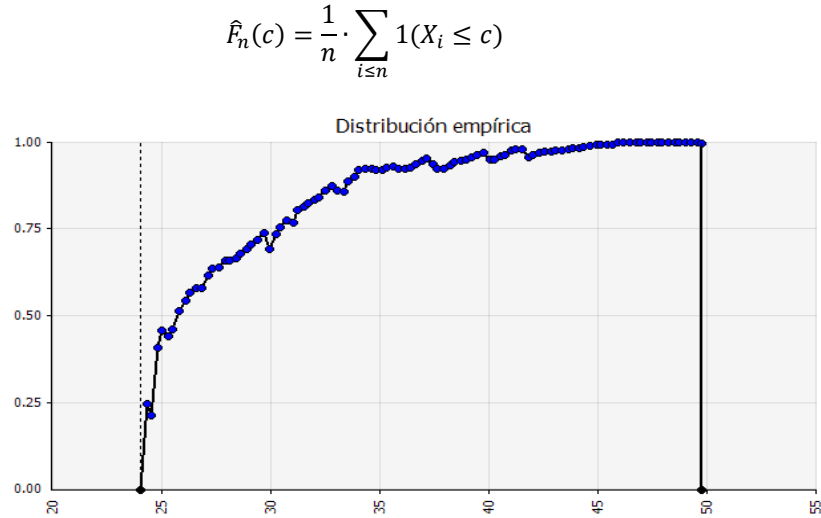
$$\widehat{W}_{\xi, \beta, \mu}(x) = 1 - \left(1 + 0.3702 \cdot \frac{x - 6.4952}{0.2718} \right)^{\frac{-1}{0.3702}}$$

Este modelo ajustado es una función de distribución condicionada a que los siniestros excedan 525.37 u.m., por tanto, esta función de probabilidades acumuladas truncada en 6.26 (logaritmo natural de 525.37), calcula la probabilidad de que un siniestro sea inferior a un valor x , una vez que haya excedido el umbral fijo en 525.37 u.m.

A su vez, la probabilidad de que sea superior a dicho valor x , vendrá dada por su función de supervivencia:

$$1 - \widehat{W}_{\xi, \beta, \mu}(x)$$

Comparando esta función de distribución teórica estimada con la función de distribución muestral $\hat{F}_n(c)$ de los 826 valores, se observa la bondad del ajuste consiguiendo el gráfico de la distribución empírica a través de la representación del par de valores $(c, \hat{F}(c))$.



A partir del modelo ajustado se pueden presentar algunos resultados:

x u.m.	$\ln(x)$	$W(\ln(x))$	$1 - W(\ln(x))$
700	6.55	17.99%	82.01%
1,050	6.96	73.42%	26.58%
1,750	7.47	89.80%	10.20%
2,450	7.80	93.67%	6.33%

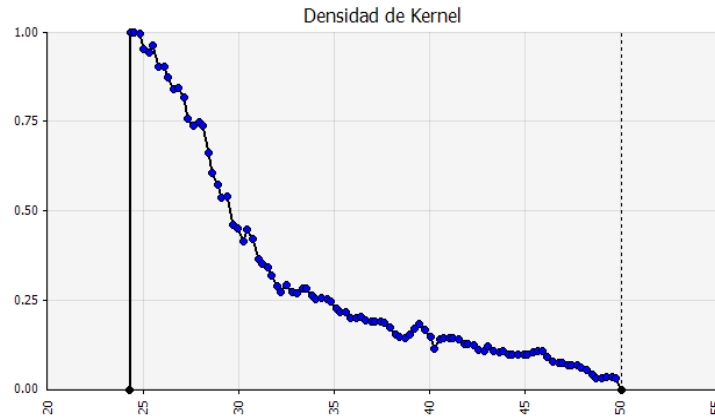
Por ejemplo, condicionado a que un siniestro supere el umbral de 525.37 u.m. la probabilidad de que el importe total del mismo sea inferior a 1,050 u.m. es de 73.42% y por tanto, la probabilidad de que sea superior a 1,050 u.m. es de 26.58%.

Luego es de esperar que, de entre los próximos 700 siniestros que superen el umbral, 73 tengan un importe inferior a 1,050 u.m. pero que 27 de ellos superen 1,050 u.m.

Derivando la función de probabilidad acumulada, la función de densidad estimada es:

$$\hat{w}_{\xi,\beta,\mu}(x) = \frac{1}{0.2718} \cdot \left(1 + 0.3702 \cdot \frac{x - 6.4952}{0.2718}\right)^{1 - \frac{-1}{0.3702}}$$

Representando esta función teórica frente a las densidades de Kernel (Anexo IV) se comprueba la bondad del ajuste



Y obteniendo su función inversa, surge la función de cuantiles teórica estimada, que indica cuál es la cuantía esperada de siniestralidad con probabilidad p :

$$\hat{W}^{-1}_{\xi,\beta,\mu}(p) = 6.4952 + \frac{0.2718}{0.3702} \cdot ((1 - p)^{-0.3702} - 1) = x_p$$

Cualquier estimación realizada a partir de los parámetros estimados deberá ser transformada exponencialmente para conocer el valor del siniestro. Algunos de los resultados obtenidos se presentan en la siguiente tabla:

p	<i>Periodo de retorno</i> $\frac{1}{1-p}$	<i>Nivel de retorno</i> (x_p) en Ln	<i>Nivel de retorno</i> (x_p) en u. m.
90%	10 <i>siniestros</i>	7.48	1,777.34
95%	20 <i>siniestros</i>	7.99	2,941.25
99%	100 <i>siniestros</i>	9.80	18,022.69
99.9%	1000 <i>siniestros</i>	15.23	4,123,972.27

Por tanto, se espera que habiendo superado un siniestro la prioridad de 525.37 u.m., con una probabilidad del 99% el importe total del mismo será inferior a 18,023 u.m. Por el contrario, con

una probabilidad del 1% (1 de cada 100 siniestros que superen el umbral) el costo total será superior a esos 18,023 u.m.

Inferencia y predicción no condicionadas:

La probabilidad absoluta de superar una cuantía extrema es una probabilidad conjunta producto de la probabilidad condicionada y de la probabilidad marginal.

Siendo p la probabilidad condicionada una vez superado el umbral u y aproximada a través del ajuste de la distribución de Pareto Generalizada, el siguiente paso es estimar la probabilidad p' no condicionada o conjunta. El objeto de conocer tal información es la toma de decisiones óptimas en cuanto a nivel de fondos no comprometidos, márgenes de solvencia o cesión al reaseguro.

En términos no paramétricos, la probabilidad marginal de que un siniestro supere 525.37 u.m. puede determinarse con la distribución empírica a partir del cociente entre el número de siniestros que han superado 525.37 u.m. (826 siniestros en 10 años) y el total de siniestros ocurridos en la cartera bajo estudio (338,128 en 10 años).

De este modo, la probabilidad conjunta de que un siniestro, de entre los próximos 1000 siniestros, supere, por ejemplo 2,450 u.m. vendrá dada por el producto de la probabilidad marginal y la probabilidad condicionada.

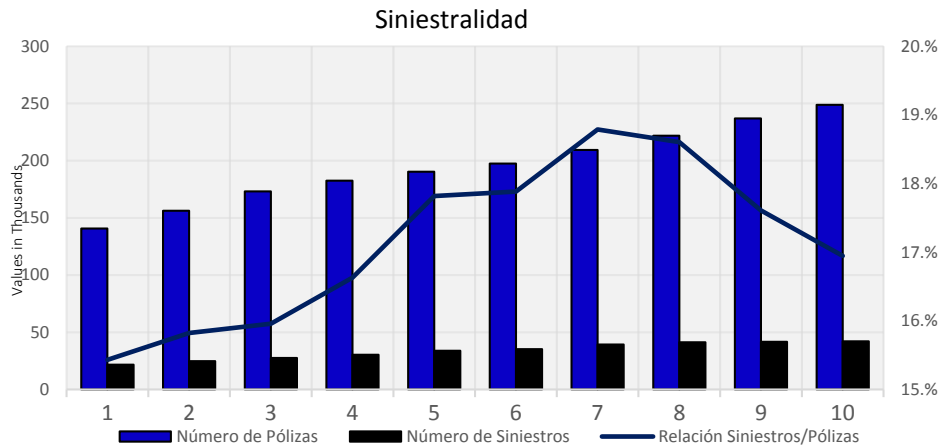
$$\text{Probabilidad marginal: } P(x > 525.37) = \frac{826}{338,128} = 0.0025$$

$$\text{Probabilidad condicionada: } P(x > 2,450 \mid x > 525.37) = 0.0105$$

$$\begin{aligned} &\text{Probabilidad conjunta de superar 2,450 u.m. en los próximos 1000 siniestros:} \\ &1000 \cdot P(x > 2,450) = P(x > 525.37) \cdot P(x > 2,450 \mid x > 525.37) = 0.0256 \end{aligned}$$

Para obtener el número de siniestros que se espera que superen el umbral de 525.37 u.m. durante el siguiente año se puede proyectar el número de siniestros, la frecuencia sobre las pólizas, o bien asumir una distribución de Poisson.

Si se observa el comportamiento de la siniestralidad en la entidad A, respecto al número de pólizas:



El eje de abscisas marca los años estudiados. El eje de ordenadas izquierdo representa el número de pólizas y de siniestros y el eje de ordenadas derecho representa la frecuencia de siniestralidad sobre pólizas.

La frecuencia de siniestralidad sobre el número de pólizas en cartera de los 10 años de estudio es muy estable, obteniendo una frecuencia alrededor del 15% y 19%.

- i) Proyectando el número de siniestros para el próximo año cabría esperar entre 43,750 y 46,760 siniestros ($R^2=96.86\%$) de los cuales, se puede pensar que entre 107 y 114 de ellos superarán el umbral de 525.37 u.m.

$$E(n_1) = \frac{826}{338,128} \cdot 43,750 = 106.88 \quad E(n_2) = \frac{826}{338,128} \cdot 46,760 = 114.23$$

- ii) Proyectando el número de pólizas ($R^2=97.92\%$) y usando la frecuencia de siniestralidad media (media aritmética ponderada por el número de pólizas), el número de siniestros esperados para el año siguiente es de 44,730 quedando dentro del intervalo anteriormente establecido.
- iii) Asumiendo una distribución de Poisson, siendo 826 el número de siniestros históricos que han superado 525.37 u.m. se asume una distribución de Poisson con parámetro $\lambda = \frac{826}{10} = 82.6 \approx 83$.

Años de estudio	Número de siniestros que han superado 525.37 u. m.
1	70
2	84
3	63
4	97
5	84
6	112
7	98
8	56
9	49
10	113

Bajo esta distribución, el número esperado de siniestros que superará el umbral el próximo año, para un 95% de nivel de confianza, será $E(n) = 110$ por tanto,

$$E(n) \cdot P(x > X \mid x > 525.37) = P(x > X) = p' = 0.189$$

Donde p' es la probabilidad conjunta de superar, por ejemplo, $X = 2,450$ u.m. en el próximo año por lo que se espera que de los siguientes 95 siniestros que superen 525.37 u.m., uno de ellos superará 2,450 u.m.

Cuantía esperada de un siniestro extremo, dada una determinada probabilidad.

Desde un punto de vista inverso, la entidad puede conocer el nivel de pérdida esperado dada una determinada probabilidad p' .

$$\hat{W}^{-1}_{\xi, \beta, \mu}(p) = 6.4952 + \frac{0.2718}{0.3702} \cdot ((1 - p)^{-0.3702} - 1) = x_p$$

Y bajo la hipótesis de que $E(n) = 110$, en base a la distribución de Poisson, la cuantía X esperada para cada p' se presenta en la siguiente tabla:

p'	Periodo de retorno $\frac{1}{p}$	X
$p' = 5.0\%$	20 años	5,887 u. m.
$p' = 2.5\%$	40 años	10,577 u. m.
$p' = 1.0\%$	100 años	27,223 u. m.
$p' = 0.5\%$	200 años	65,058 u. m.

Por lo que con una probabilidad de 0.5% el próximo año no se superará la cuantía 65,058 u.m. y el periodo de retorno de un siniestro de estas características será de 200 años.

3.3 Frecuencia de valores extremos por el método de Gumbel.

Este método no paramétrico busca la probabilidad de que el k –ésimo mayor siniestro de entre n observaciones se exceda h veces en los N ensayos próximos.

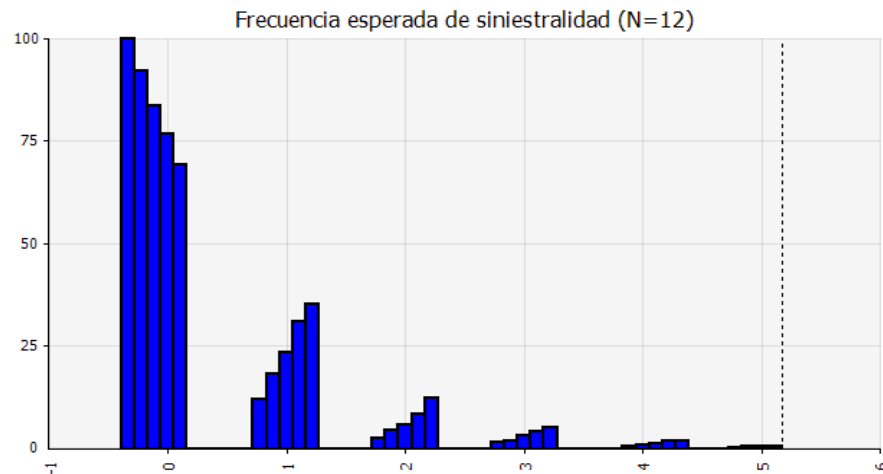
En este trabajo de investigación, la entidad A ha cedido 10 años de observación y tomando los máximos mensuales, se cuenta con una muestra histórica de $n = 12$.

El método de Gumbel permite obtener las probabilidades de frecuencia de ocurrencia para los próximos, por ejemplo, 1, 2, y 3 años ($N = 12, N = 24$ y $N = 36$).

En los 10 años estudiados, los cinco siniestros máximos registrados fueron:

<i>Siniestro de orden k</i>	<i>Costo del siniestro u. m.</i>
1	2,370,186
2	1,979,383
3	1,371,797
4	1,197,735
5	1,187,725

El siguiente gráfico muestra las frecuencias esperadas para los próximos 12 máximos mensuales. En el eje de abscisas se muestra el número h de siniestros esperados y en el eje de ordenadas la frecuencia esperada de que ese hecho ocurra, $w(h)$;

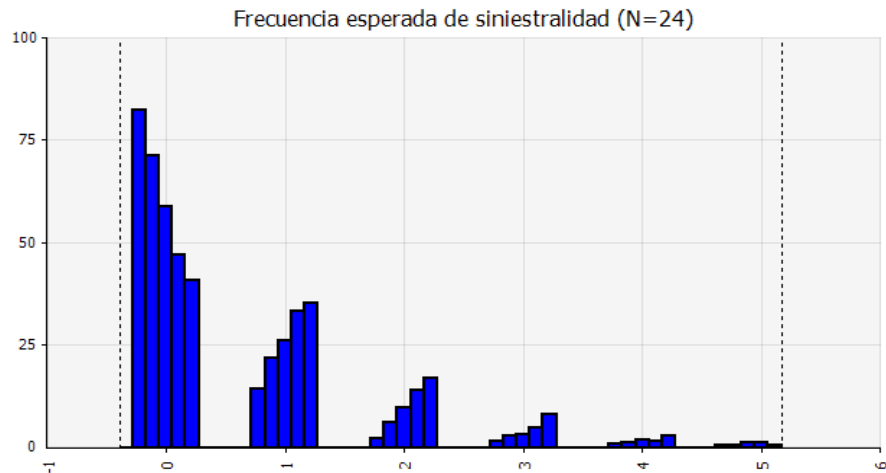


A través de este método se estima que, durante el próximo año (12 observaciones máximas mensuales), el máximo valor mensual histórico $k = 1$ no sea superado en el año siguiente $h = 0$

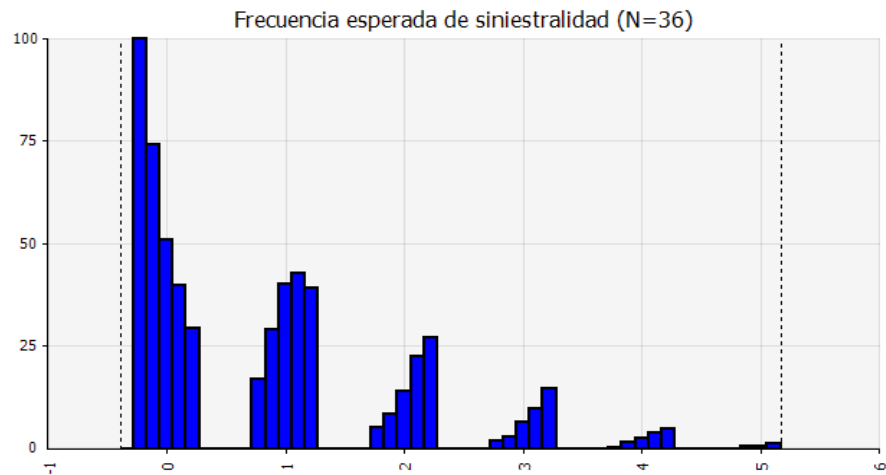
con una probabilidad del 91%. Mientras que a medida que aumentamos el valor de k , bajando posiciones en la jerarquía descendente, la probabilidad de que dicho máximo no sea superado se va reduciendo.

Por el contrario, se estima que el mayor máximo histórico sea superado 1 vez $h = 1$, es decir, uno de los doce próximos máximos mensuales será superior al mayor valor histórico con una probabilidad del 8%. A medida que se toma una posición inferior en la jerarquía, la probabilidad incrementa, siendo 29% la probabilidad de que el 5º máximo sea superado 1 vez entre los 12 próximos máximos. Las mayores frecuencias estimadas, considerando los 5 valores de k , se centran en $h = 0$ por lo que es más probable que no se supere ninguno de los 5 extremos históricos en ninguna de las 12 observaciones máximas futuras.

De manera similar se puede estimar la probabilidad de que los máximos históricos sean superados en alguna de las $N = 24$ y $N = 36$ siguientes máximas observaciones:



La probabilidad de que no se supere ($h=0$) el mayor de los máximos mensuales históricos es superior al 80%, la probabilidad de que dicho máximo sea superado alguna vez es inferior al 20%. Tomando $k=1$, la probabilidad acumulada alcanza el $W(h) = 99\%$ de confianza en $h=2$, es decir que existe una probabilidad de 1% de que el máximo histórico sea superado 3 o más veces.



El cuarto mayor valor ($k=4$) presenta una probabilidad de 32.8% de ser superado en una ocasión y de no ser superado en ninguna de 34.6%. Para el quinto mayor siniestro mensual histórico la probabilidad de no ser superado en los próximos tres años se ha reducido a 26.4%, con lo cual existe una probabilidad de 83.6% de que el quinto máximo histórico se supere 1 o más veces, como máximo mensual entre los próximos 36 máximos mensuales.

Continuando para sucesivos años, a medida que se aumenta el número N de observaciones futuras, las probabilidades de que los máximos no sean superados comienzan a disminuir y a incrementar las probabilidades de superar los máximos mensuales históricos.

Conclusiones.

La presente investigación actuarial ha tratado los grandes riesgos por ser aquellos que, a pesar de su baja frecuencia, pueden ocasionar consecuencias devastadoras.

Exponer la utilidad de la Teoría del Valor Extremo como medio para adoptar decisiones financieras óptimas en relación con los siniestros esperados más graves, complementan la teoría clásica de riesgo y permite obtener resultados válidos para la toma de decisiones, de manera que se garantice la solvencia y la capacidad de afrontar situaciones extremas futuras, mediante el uso de capitales propios o el reaseguro.

La justificación de este trabajo se deriva de los permanentes cambios que la industria financiera y el sector asegurador están sufriendo. Debido al avance tecnológico y al desarrollo económico, existe mayor exposición a siniestros con consecuencias de alto nivel que pueden llegar incluso a ser catastróficas. Se buscó obtener un equilibrio de solvencia para que las primas cobradas por anticipado se adecuen a la suma de las contraprestaciones pagadas, a fin de mantener la estabilidad de las entidades aseguradoras.

El seguro y el reaseguro, encargados de proporcionar cobertura ante la ocurrencia de sucesos económicamente desfavorables, aceptan el riesgo de situaciones que pueden dañar su estabilidad. La principal consecuencia es la preocupación por la solvencia ante quien deposita su confianza en ellas y la consecuente necesidad de coberturas que garanticen la supervivencia de la misma.

Es aquí donde la Teoría del Valor Extremo comienza a jugar un papel importante, a partir de modelos que cuantifican valores extremos se puede medir su frecuencia y consecuencias de un modo estadísticamente óptimo.

El actual marco actuarial se basa en el proyecto de Solvencia II, justificando este trabajo de investigación. Todos los documentos de trabajo revisados hacen continuas alusiones a los eventos extremos y a su relación con la valoración de la solvencia. Se desprende del proyecto de Solvencia II la necesidad de tener en cuenta las distribuciones de probabilidad que describen las colas, ya que son dichos sucesos los que en mayor medida pueden comprometer la posición financiera de una entidad.

Las asunciones de la teoría del riesgo (basadas en el teorema de límite central y, por tanto, en el comportamiento medio de las observaciones) desprecian la información de sucesos que presentan una probabilidad de acaecimiento extremadamente pequeña con consecuencias graves en términos monetarios, puesto que contaminan el comportamiento central e incrementan la dispersión.

El análisis de una cartera de siniestros no se debe limitar a realizar una transformación logarítmica que permita estandarizar el estudio. Las distribuciones de cola gruesa muestran mayor masa de probabilidad en la cola que distribuciones como la normal.

El futuro es la inferencia basada en mezclas de distribuciones que ajusten la masa por un lado y separadamente las colas a través de la teoría que las explica. Bajo el teorema de Fisher – Tippet, se puede aproximar la distribución de probabilidad de máximos mediante una de las tres distribuciones del valor extremo, cuya forma estándar es la Distribución Generalizada del Valor Extremo. El parámetro de la distribución, ξ , indica el grueso de la cola de la distribución, que será suave o finita cuando el parámetro es negativo y de cola media o gruesa cuando el parámetro es nulo o positivo. Cuanto mayor sea ξ , más gruesa es la cola. Por otro lado, el teorema de Pickands – Balkema – de Haan, establece que la distribución generalizada de Pareto es la adecuada cuando se pretenden modelar los extremos que exceden un cierto umbral u . El interés por esta distribución es inherente al proceso actuarial y al estudio de la siniestralidad por encima de un valor elevado, haciendo referencia inmediata al reaseguro de exceso de pérdida.

Con el objeto de manifestar la utilidad de esta teoría en la gestión de riesgos de las entidades aseguradoras, se ha realizado una aplicación a la cartera de siniestros de una entidad aseguradora representativa dentro del sector automotriz.

Los principales resultados son:

- La entidad es estable y no presenta tendencias.
- No se puede negar ni obviar que existe recurrencia en la concentración de siniestros graves en determinados periodos del año, sin embargo, los test realizados contrastan que dichas diferencias no son significativas.
- No se ha detectado relación entre el coste de la siniestralidad grave y el tiempo de declaración.
- Se ha conseguido modelar las colas de las distribuciones de siniestralidad de ambas entidades aseguradoras. La Distribución Generalizada de Pareto permite realizar inferencia y, por ende, tomar decisiones a partir de la información que ofrece el modelo el cual permite estimar; probabilidades y cuantiles, esperanzas condicionadas, un nivel de reservas adecuado y medidas de riesgo como VaR y TVaR
- Los modelos de la Teoría Clásica son válidos para modelar la masa pero no son adecuados para extrapolar el comportamiento de la cola. Su empleo, basado en la media y la

dispersión, genera una mala valoración de los altos cuantiles y probabilidades de siniestros graves.

La investigación actuarial debe apostar por líneas futuras que aporten valor al conocimiento existente y su aplicación a un sector cada vez más innovador. La Teoría del Valor Extremo es un campo que necesita ser explorado, tanto desde el punto de vista teórico como práctico.

El estudio independiente de los extremos no sustituye ni reemplaza el proceso de riesgo según la teoría clásica, sino que la complementa.

Bibliografía.

- [1] Alegre, A.; Sarrasi, J., 1995. *Modalidades alternativas de reaseguro basadas en el orden de riesgos*.
- [2] Beard, R.E.; Pentikainen, T.; Pesonen, E., 1984. *Risk Theory*; Chapman & Hall, Londres.
- [3] Beirlant, J.; Teugels, J.L.; Vynckier, P., 1996. *Practical Analysis of Extreme Values*; Leuven University Press. Leuven.
- [4] Bensalah, Y., 2000. *Steps in applying Extreme Value Theory to finance: a Review*, Bank of Canada Working Paper 2000-20.
- [5] Bühlmann, H., 1970. *Mathematical Methods in Risk Theory*; Springer Verlag, Heidelberg
- [6] Embrechts, P.; Klüppelberg, C.; Mikosch, T., 1997: *Modelling extremal events for Insurance and Finance; Applications of Mathematics*. Springer-Verlag Berlin Heidelberg.
- [7] Fisher, R.A. Y Tippett, L.H.C., 1928: *Limiting Form of the Frequency Distribution of the largest or smallest member of a sample*; Proceedings of the Cambridge Philosophical Society, Vol. 24.
- [8] García Pérez, *La teoría del valor extremo: una aplicación al sector asegurador*; Universidad de Alcalá de Henares, Madrid
- [9] Gumbel, E.J., 1954. *Statistical theory of extreme values and some practical applications*; United States Department of Commerce. National Bureau of Standards, Washington.
- [10] Gumbel, E.J., 1958. *Statistics of Extremes*; Columbia University Press, New York.
- [11] Hoel Port Stone, 1972. *Introduction to Stochastic Processes*; University of California, Los Angeles.
- [12] Stuart A. Klugman, Drake; Harry H. Panjer; Gordon E. Willmont; *Loss Model From Data to Decisions*; University of Waterloo
- [13] Mcneil, A.J.; Saladin, T., 1997. *The Peaks over Thresholds Method for Estimating High Quantiles of loss Distributions*.
- [14] Panjer, H.H.; Willmot, G.E., 1992. *Insurance Risk Models. Society of Actuaries*, Schaumburg.

Anexos.

Anexo I

Proceso de recuento

Un proceso de recuento, $\{N(t), t \geq 0\}$, es una representación del número total de acontecimientos u ocurrencias hasta el momento t . Estos procesos tienen las siguientes propiedades:

- $N(t) \geq 0$.
- $N(t)$ es un valor entero.
- Si $t_0 < t_1$, entonces $N(t_0) \leq N(t_1)$. Es decir $N(t)$ es no decreciente.
- Si $t_0 < t_1$, entonces $N(t_1) - N(t_0)$ es el número de ocurrencias en $(t_0, t_1]$.

Incrementos independientes

Un proceso estocástico tiene incrementos independientes o de renovación si para algún t_i , $i = 0, \dots, n$ y $n \geq 1$ tal que $0 = t_0 < t_1 < \dots < t_n$, los incrementos $N(t_i) - N(t_{i-1})$ son mutuamente independientes. Es decir, la ocurrencia de un suceso no influye en la ocurrencia de sucesos posteriores.

Incrementos estacionarios

Un proceso estocástico tiene incrementos estacionarios si la distribución de $N(t_1) - N(t_0)$ depende únicamente de la duración del intervalo $(t_0, t_1]$ y no de la localización de éste. Si diferentes intervalos coinciden en duración, éstos tendrán la misma probabilidad.

Por lo tanto, un proceso de recuento será de incrementos independientes si el número de ocurrencias que tienen lugar en intervalos de tiempo disjuntos son variables aleatorias independientes. Y será de incrementos estacionarios si el número de ocurrencias que tienen lugar en un intervalo de tiempo es el mismo en intervalos de igual longitud. Un ejemplo de estos tipos de procesos son los de renovación, dentro de los cuales hallamos como caso particular el proceso de Poisson.

Proceso de Poisson

Un proceso de recuento, $\{N(t), t \geq 0\}$, es un proceso de Poisson homogéneo con tasa $\lambda > 0$ si se verifica:

1. $N(0) = 0$.
2. Es de incrementos independientes y estacionarios.
3. $P(N(t+h) - N(t) = 1) = \lambda h + o(h)$, siendo h y $o(h)$ un intervalo corto de tiempo tal que $\lim_{h \rightarrow 0} \frac{o(h)}{h} = 0$.
4. $P(N(t+h) - N(t) \geq 2) = o(h)$, esta propiedad elimina la probabilidad de más de un suceso en un mismo instante.

Si un proceso de recuento es un proceso de Poisson, entonces el número de siniestros $N(t)$ para $\forall t > 0$ se distribuye como una distribución *Poisson*(λt).

Tiempos de interocurrencia

Sean $0 < T_1 < T_2 < T_3 < \dots$ los momentos de ocurrencia, se definen los tiempos de interocurrencia como

$$W_i = T_i - T_{i-1} \quad i = 1, 2, \dots$$

Siendo $W_1 = T_1$

Proceso de renovación

Un proceso de recuento, $\{N(t), t \geq 0\}$, es de renovación cuando los tiempos entre ocurrencias son variables aleatorias positivas independientes e idénticamente distribuidas.

En particular, el proceso de Poisson con parámetro λ es un proceso de renovación donde los tiempos de interocurrencia son variables independientes e idénticamente distribuidas, con distribución exponencial de media $\frac{1}{\lambda}$.

Proceso de Poisson compuesto

Si se combina un proceso de Poisson $\{N(t), t \geq 0\}$ con una sucesión de variables aleatorias $\{X_i\}_{i=1}^{\infty}$, independientes e idénticamente distribuidas, e independientes $N(t)$, se obtiene un proceso de Poisson compuesto $\{S(t), t \geq 0\}$:

$$S(t) = \sum_{i=1}^{N(t)} X_i$$

Si $N(t) = 0$ entonces $S(t) = 0$

La manera de interpretar el proceso anterior en una cartera de seguros no vida es considerar $N(t)$ como el número de siniestros ocurridos hasta el momento t y X_i como la cuantía individual del siniestro i -ésimo. Este tipo de proceso, que recoge la siniestralidad agregada de la cartera, forma parte del proceso de las reservas.

Anexo II

Distribución de Pareto

Esta es una de las distribuciones frecuentemente empleadas como distribución analítica de la cuantía de siniestralidad, sin embargo, esta distribución no proporciona buenos ajustes para el ajuste de cola de la distribución.

La distribución de Pareto presenta la siguiente función de distribución:

$$F(x) = P(S \leq x) = 1 - \left(\frac{c}{x}\right)^b$$

Donde S representa la variable aleatoria de la cuantía de siniestralidad esperada. Su función de densidad se expresa como:

$$f(x) = \frac{b \cdot c^b}{x^{b+1}}$$

$$\text{Con media } \int_0^x x \cdot dF(x) = \frac{b \cdot c}{b-1} \quad \text{y varianza } \frac{b \cdot c^2}{(b-2) \cdot (b-1)^2}$$

Si $c > X$, la función toma valores negativos, y es por ello que esta distribución sólo está definida para $X \geq c$, esto es, para valores de siniestralidad que superen una prioridad c .

Se exige que b sea un parámetro positivo e indica el grueso de cola de la distribución, cuanto más pequeño sea b , más fuerte es la cola.

Distribución de Poisson

La distribución de Poisson es el caso límite de la distribución binomial, cuando la probabilidad de ocurrencia de un siniestro es pequeña y el número de casos n es elevado. En el límite, la probabilidad de que ocurran exactamente k siniestros en el periodo considerado es:

$$p_k = \frac{e^{-\lambda} \cdot \lambda^k}{k!}, \quad \lambda \geq 0$$

Donde λ es el parámetro característico de la distribución Poisson y a su vez la media y varianza de la misma.

La distribución de Poisson es fundamental como modelo para medir la ocurrencia de eventos aleatoriamente en el tiempo, para esto;

- El número de siniestros que ocurren en dos intervalos disjuntos de tiempo son independientes.
- Se excluye la multisiniestralidad en el periodo de exposición al riesgo.
- La probabilidad de que el siniestro ocurra en un momento fijo del tiempo es cero, se excluyen momentos especiales de ocurrencia.
- El número esperado de siniestros es una función continua de t .
- Ausencia de estacionalidad.

Distribución de Poisson Mixta

La distribución de Poisson no resulta del todo válida en la práctica, debido a las restricciones que presenta. La distribución de Poisson mixta añade una aleatoriedad en el parámetro característico λ , dejando de ser determinista para pasar a ser aleatorio.

Se presta atención a la independencia en la ocurrencia de siniestros. La presencia de factores que puedan provocar siniestros correlacionados no se contempla en la distribución de Poisson, por ello, en la distribución mixta se introduce una variable auxiliar o proceso que controle los cambios en la propensión al riesgo.

La probabilidad de que ocurran exactamente k siniestros en el periodo de tiempo dado es:

$$p_k = \int_0^{\infty} \frac{e^{-\lambda\theta} \cdot (\theta\lambda)^k}{k!} \cdot dF(\theta)$$

Donde $F(\theta)$ es la función de distribución del factor multiplicativo θ , denominada variable mixta (*mixing*), con $E[\theta] = 1$

Si la intensidad del siniestro durante el periodo de tiempo es el nivel esperado, entonces $\theta = 1$, por el contrario si la intensidad es mayor que la esperanza $\theta > 1$ y si fuera menor $0 < \theta < 1$.

Si la ponderación θ es fija, se sigue manteniendo la hipótesis de independencia y estaríamos ante un proceso de Poisson con parámetro $\lambda\theta$.

La media de la variable aleatoria k que se distribuye como una Poisson mixta, coincide con la media del parámetro de Poisson que varía aleatoriamente:

$$E(k) = E(E(k|\theta)) = E(\lambda\theta) = \lambda \cdot E(\theta) = \lambda \cdot 1 = \lambda$$

Anexo III

Teorema de Pickands-Balkema-de Haan

Sea $X_{1,n} \geq X_{2,n} \geq \dots \geq X_{k,n} \geq \dots \geq X_{n,n}$ un conjunto de variables aleatorias independientes e idénticamente distribuidas con una función de distribución común $F(x) = P(X_{i,n} \leq x)$ y donde M_n representa el máximo de todas ellas, $M_n = X_{1,n} = \max(X_{1,n}, X_{2,n}, \dots, X_{k,n}, \dots, X_{n,n})$, en el límite se verifica que:

$$P\left(\frac{M_n - a_n}{b_n} \leq x\right) \rightarrow G(x) \text{ cuando } n \rightarrow \infty$$

Donde

$$G_\xi(x) = \exp\left\{-\left(1 + \xi \cdot \frac{x - \mu}{\sigma}\right)^{-\left(\frac{1}{\xi}\right)}\right\}$$

Entonces, para valores suficientemente grandes de la prioridad u ,

$$F^u(x) = P(X - u \leq y \mid X > u) \sim W_\xi(y) \quad \text{para } 0 \leq y \leq x_0 - u$$

Demostración:

$$\ln G_\xi(x) = -\left(1 + \xi \cdot \frac{x - \mu}{\sigma}\right)^{-\left(\frac{1}{\xi}\right)}$$

Para valores altos de x , la aproximación de Taylor indica que:

$$\ln F(x) \approx -[1 - F(x)]$$

Aplicando esta aproximación a la expresión anterior:

$$-[1 - G_\xi(x)] = -\left(1 + \xi \cdot \frac{x - \mu}{\sigma}\right)^{-\left(\frac{1}{\xi}\right)}$$

$$G_\xi(x) = 1 - \left(1 + \xi \cdot \frac{x - \mu}{\sigma}\right)^{-\left(\frac{1}{\xi}\right)}$$

$$\text{Si } W_\xi(y) \sim P(X - u \leq y \mid X > u) \quad y \quad P(X - u \leq y \mid X > u) = \frac{F(x) - F(u)}{1 - F(u)}$$

Entonces
$$W_{\xi}(y) \sim \frac{G_{\xi}(x) - G_{\xi}(u)}{1 - G_{\xi}(u)} = \frac{-\left(1 + \xi \cdot \frac{x - \mu}{\sigma}\right)^{-\left(\frac{1}{\xi}\right)} + \left(1 + \xi \cdot \frac{u - \mu}{\sigma}\right)^{-\left(\frac{1}{\xi}\right)}}{\left(1 + \xi \cdot \frac{u - \mu}{\sigma}\right)^{-\left(\frac{1}{\xi}\right)}} = 1 -$$

$$\left(\frac{\sigma + \xi \cdot (x - \mu)}{\sigma + \xi \cdot (u - \mu)}\right)^{-\left(\frac{1}{\xi}\right)}$$

Siendo $x = y + u$ y $\beta = \sigma + \xi \cdot (u - \mu)$

$$W_{\xi}(y) = 1 - \left(\frac{\beta + \xi \cdot y}{\beta}\right)^{-\left(\frac{1}{\xi}\right)} = 1 - \left(1 + \frac{\xi \cdot y}{\beta}\right)^{-\left(\frac{1}{\xi}\right)}.$$

Anexo IV

Densidad de Kernel

Sea

$$g_b(x, x_i) = \frac{1}{n} \cdot \frac{k}{b} \cdot \left(\frac{x - x_i}{b} \right) \quad b > 0$$

Donde b es el ancho de banda elegido y k es la función de Kernel tal que $\int k(y) dy = 1$

Si k es no negativa, entonces $\frac{k}{b} \cdot \left(\frac{x - x_i}{b} \right)$ es una densidad con parámetros de localización y escala, x_i y b respectivamente.

Sumando todos los términos individuales se obtienen las densidades de Kernel.

$$f_{n,b}(x, x_i) = \sum_{i=1}^n g_b(x, x_i) = \frac{1}{n \cdot b} \cdot \sum_{i=1}^n k \cdot \left(\frac{x - x_i}{b} \right) \quad b > 0$$

Si el ancho de banda b o parámetro de escala es pequeño, todavía se pueden reconocer términos de $g_b(x, x_i)$ representando el comportamiento real de los datos. Si por el contrario, para valores muy grandes de b , se produce una sobre suavización que puede ocultar la detección de variaciones claves que proporcionan los datos.