

1. INFORME EJECUTIVO

El presente informe tiene por objetivo realizar un análisis que analice las pérdidas de la empresa por unidad de negocio, la cual se extenderá al resto de unidades de negocio, es decir, a las distintas especies piscícolas. Para ello veremos en primer lugar cómo se distribuyen los datos, los cuales se comportan como una distribución asimétrica tanto en severidad como en frecuencia, tanto semanal como mensual. Las medidas de adecuación que utilizaremos serán mediante una distribución Poisson y una de Binomial Negativa. Nos quedamos con el enfoque utilizado a través de esta última por estar más ajustada y arrojar un menor AIC y BIC. Seguidamente realizaremos un análisis de los valores extremos a través de la Distribución Generalizada de Pareto observando que el mejor ajuste lo proporciona la distribución Burr. Adicionalmente realizando el método de la teoría de los valores extremos, mediante el POT y el GEV, este último realizado mediante el ajuste de Block Máxima observaremos que el impacto producido por las pérdidas teóricas es capaz de explicar las pérdidas incurridas. Realizaremos a su vez el análisis de las pérdidas agregadas observando que la mejor aproximación hacia la Distribución de Pérdidas Agregadas la ofrece la Distribución Normal-Power y, por otra parte, a través del método Pareto en donde calcularemos la función de distribución discretizándola. Finalmente calcularemos el VaR para todos y cada uno de los supuestos que se exponen para el 95 y 99% de intervalo de confianza asumiendo diferentes escenarios y con distribuciones de probabilidad normales y t-Student.

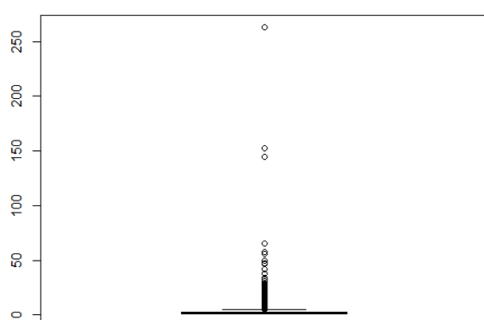
2. ANÁLISIS EXPLORATORIO DE LOS DATOS

El conjunto de datos con los que trabajaremos será el proporcionado por la compañía de Reaseguro Danesa de Copenhague y comprende 2167 observaciones. Cada una de ellas son las pérdidas por incendio durante el período 1980 a 1990. Se han ajustado a la inflación para reflejar los valores de 1985 y se expresan en millones de coronas danesas.

La media de pérdidas se sitúa en 3.38 mientras que la mediana está en 1.77, todo parece indicar que existen valores extremos. Analizando los cuartiles observamos que en el último decil existe una gran cola de valores extremos debido a la diferencia significativa de pérdidas entre el percentil 99 con un 26.04 y el percentil 100 con unas pérdidas de 263.25.

A través de un boxplot se observa la asimetría en nuestros datos, lo cual parece indicar la presencia de valores extremos en nuestra distribución. Según el coeficiente de asimetría se observa la asimetría hacia la derecha, siendo este valor de 18.74. Indicar que si este coeficiente fuese 0 nos encontraríamos ante la simetría de la distribución en los datos comportándose como una

Distribución Normal. Además, la Kurtosis es bastante acusada, comportándose como una Distribución Leptocúrtica con un valor aproximado de 485.64. Indicar a su vez que valores superiores de 3 indican la ausencia de una Distribución Normal en nuestros datos.



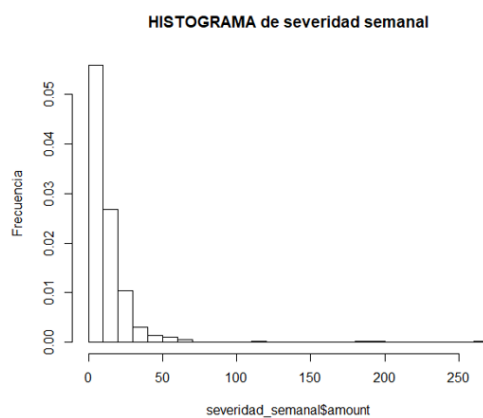
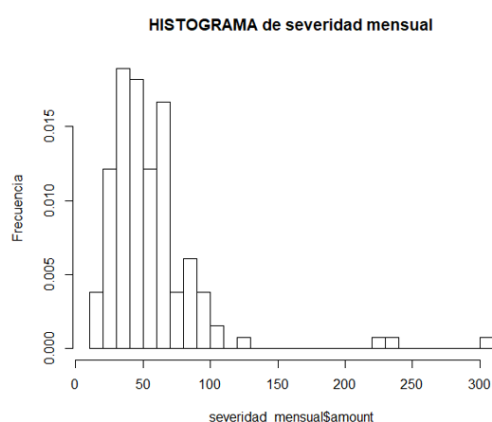
Valores muy concentrados en el primer cuartil y valores extremos muy dispersos en el último. Existe una asimetría significativa en los datos

3. SELECCIÓN DEL MODELO: INFERENCIA PARAMÉTRICA

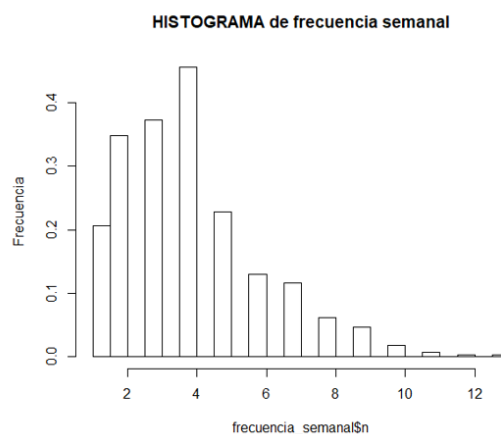
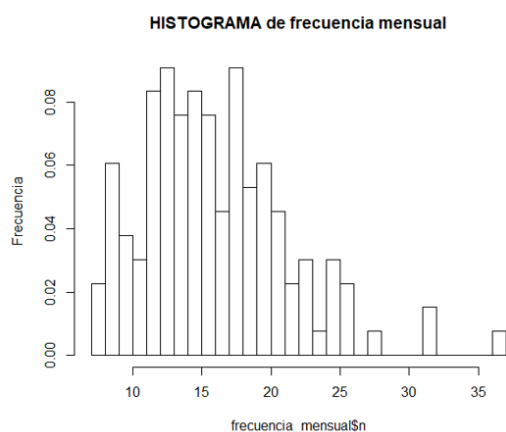
a. Ajuste de distintas distribuciones

En primer lugar, tendremos que discernir entre severidad y frecuencia para poder realizar posteriormente nuestro análisis estadístico y poder extraer conclusiones significativas por riesgo operativo, destacando el período de tiempo en donde se produce un mayor impacto de pérdidas para la compañía. Para ello, agruparemos las severidades y las frecuencias por franjas temporales de semanas y meses.

Por una parte, realizando un análisis de las severidades mensuales se observa que la proporción de pérdidas se concentra dentro del intervalo 40 a 75 millones de coronas danesas. En cuanto a las severidades semanales observamos en cuanto a frecuencia que la mayor parte de las pérdidas se producen dentro del intervalo de 0 a 40, aproximadamente. El gráfico que representa el análisis por severidades tanto mensuales como semanales serían los siguientes:



Por otra parte, realizando un análisis de las frecuencias mensuales se observa que desde el mes 10 al mes 20¹ existe una mayor concentración de pérdidas. Realizando el mismo análisis para las semanas se observa como las primeras semanas hasta la semana 4 existe una mayor frecuencia de pérdidas ocasionadas para la empresa. El gráfico que representa el análisis por frecuencia tanto mensuales como semanales serían los siguientes:



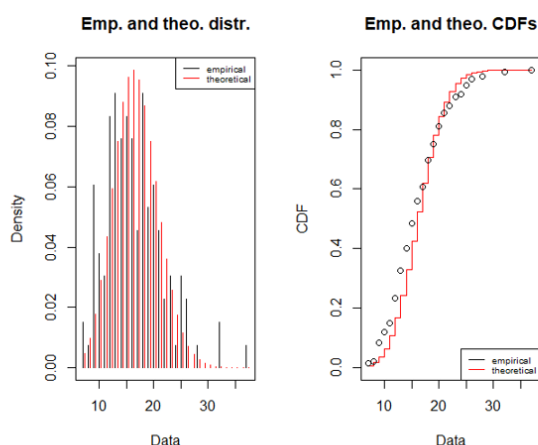
b. Medidas de adecuación

¹ En el archivo. R que se adjunta se detalla la explicación, así como la representación gráfica derivado de este análisis.

Tendremos que realizar un ajuste de máxima verosimilitud a través de la distribución Poisson, para ello realizaremos tal análisis diferenciando si la frecuencia del evento ha sido mensual o semanal. Hay que destacar que la lambda es la media teniendo en cuenta que el lambda es $n \cdot p$.

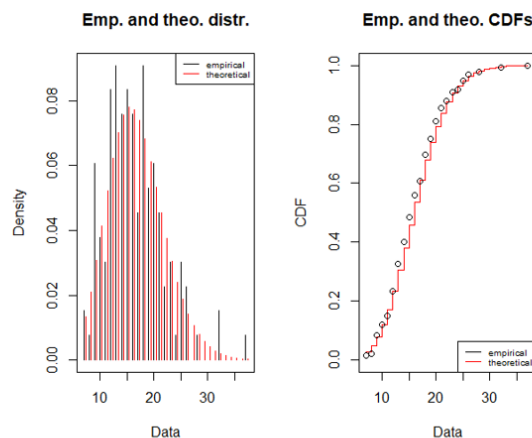
Además, también tendremos que realizar un análisis según la binomial negativa comprobando las principales diferencias entre cada una de ellas. Todos los gráficos que se exponen a continuación son la función de densidad por una parte y por otra parte la función de distribución. Vemos que tanto lo empírico como lo teórico está bastante bien ajustado, sobre todo si la frecuencia de nuestra distribución es mensual.

DISTRIBUCIÓN DE POISSON



```
Parameters:
  estimate Std. Error
lambda 16.41667 0.3526594
```

DISTRIBUCIÓN BINOMIAL NEGATIVA



```
Parameters:
  estimate Std. Error
size 25.33576 7.811960
mu 16.41691 0.452729
```

c. Contraste de bondad del ajuste

El estadístico de contraste es la Chi-Cuadrado, ajustando los valores al modelo realizando un contraste de hipótesis en donde la hipótesis nula es la normalidad de la distribución. Podemos sacar como conclusión que existe un mejor ajuste para la Binomial Negativa que para la de Poisson ya que el AIC y BIC son menores.

Los Chi-breaks que establezco serán de 0:4, 9 ya que considero que son los valores que mejor definen los valores que tenemos de la muestra².

² Este criterio resultará fundamental para calcular el estadístico de la Chi-cuadrado

A su vez existe un mejor ajuste cuando realizamos el análisis agrupando por frecuencias mensuales que semanales ya que en base a los mismos criterios el AIC de la Binomial Negativa para la frecuencia semanal es de 2330.54 y para la mensual resulta ser de 806.35.

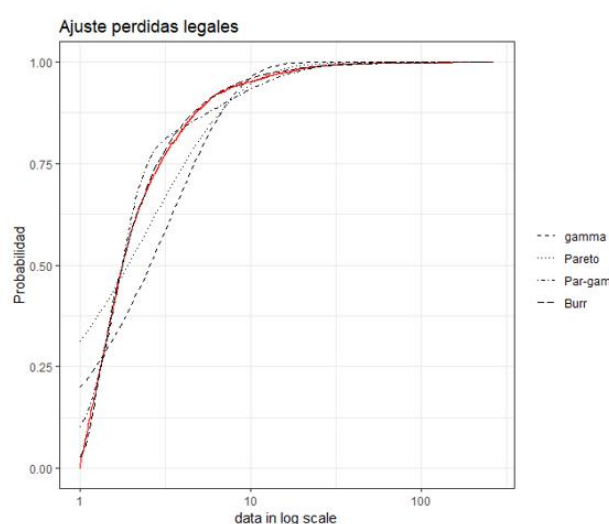
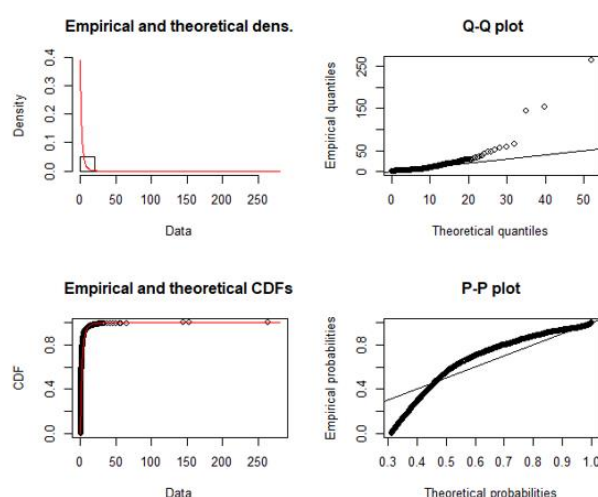
En base a todos estos criterios el mejor resultado lo ofrece la Binomial negativa agrupando por frecuencias mensuales.

4. Análisis de los valores extremos

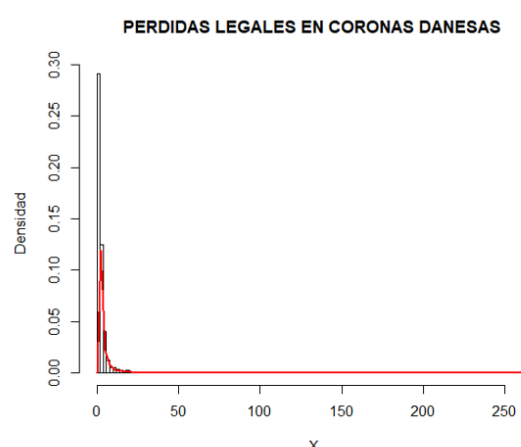
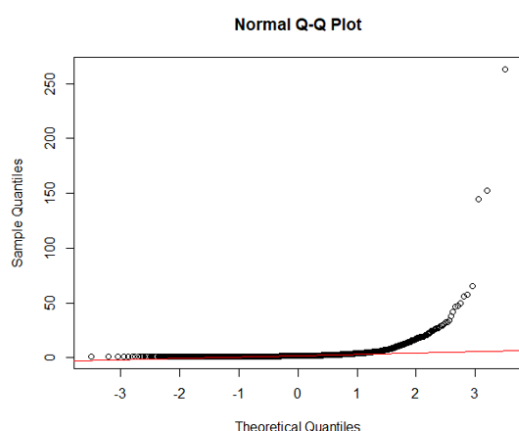
a. Distribución Generalizada de Pareto

Para realizar el análisis de la Distribución Generalizada de Pareto resulta fundamental para detectar las colas pesadas en nuestra distribución. Las distribuciones de probabilidad más ampliamente usadas para poder describir las pérdidas de nuestra empresa son la distribución de Weibull, Gamma, Humbel para extremos, distribuciones lognormal y Gamma Logarítmica. Se entiende por cola pesada aquella función de densidad que tiende a cero más lentamente que una función exponencial. El objetivo es determinar la existencia o no de colas pesadas en las funciones de distribución de probabilidad y estimar los parámetros asociados a estas distribuciones a través de la distribución generalizada de Pareto.

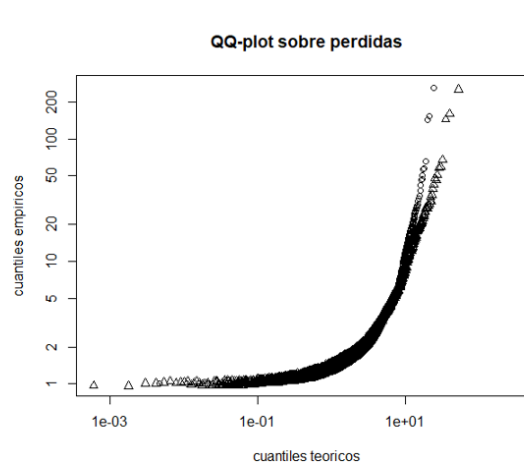
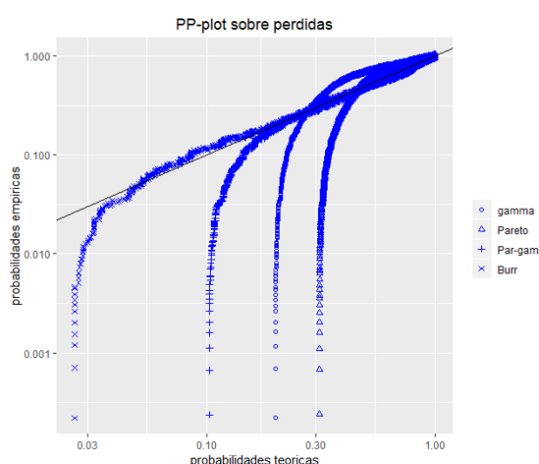
Comparando las densidades ajustadas en el gráfico de la derecha



Realizamos a continuación la selección de la distribución, de nuevo lo podemos observar en el gráfico de la derecha



Finalmente representamos las funciones de distribución, tanto la Gamma, la Pareto, la mixtura de Pareto- Gamma y la Burr. La que mejor ajusta es la distribución de Burr con lo teórico, en cambio la distribución empírica de Pareto es la que menos con respecto a la teórica.



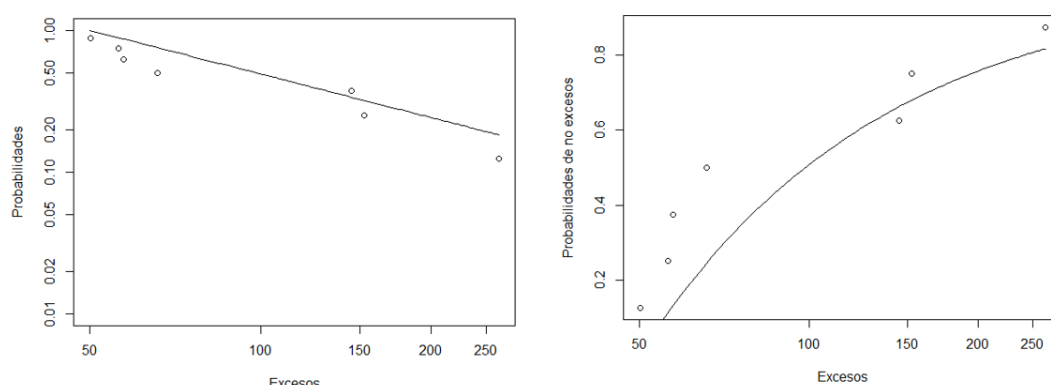
b. Distribución Generalizada de valores extremos

El método de la teoría de valores extremos (EVT) se utiliza para modelar excesos sobre un umbral.

Es importante destacar que la modelización de las colas de la distribución puede desarrollarse a través de dos diferentes estrategias. Por una parte, el llamado Block Maxima (GEV) caracterizando la distribución la cual es basa en el ajuste de la distribución de los valores máximo o mínimos y por otra parte el Peaks Over Threshold (POT) el cual ofrece un análisis de los valores extremos a partir del análisis de los valores que exceden cierto umbral.

Para realizar el POT, he establecido como umbral el valor 50, el cual ofrece una aproximación a la modelización de los eventos de la cola de la distribución. La gran mayoría de distribuciones continuas pueden ser ajustadas si se tiene en cuenta un umbral (u) suficientemente alto. Adicionalmente, el método requiere que exista una cantidad suficiente de datos que estén por encima de este caso.

Podemos visualizar las colas a través de un gráfico en donde se relacione su probabilidad con los puntos señalados como exceso. A continuación, añadimos las probabilidades teóricas de la Distribución de Pareto, determinando los parámetros necesarios para su representación y establecemos la función de distribución acumulada. Las dos gráficas representan los excesos, en donde la primera relaciona el exceso de potenciales pérdidas con las probabilidades reales y la segunda relaciona el exceso de las pérdidas con las probabilidades de los no excesos en nuestra distribución de datos.



Por otra parte, la distribución de los valores extremos generalizados (GEV) se ajusta mediante Block Máxima, la cual se basa en el ajuste de la distribución de valores máximos o mínimos. Extraemos las pérdidas máximas ocasionadas cada año y obtenemos los cuatro primeros valores de la variable. Realizamos la misma operativa pero esta vez con el modelo Poisson- Generalizada de Pareto y obtenemos las siguientes diferencias:

MODELO BLOCK MÁXIMA

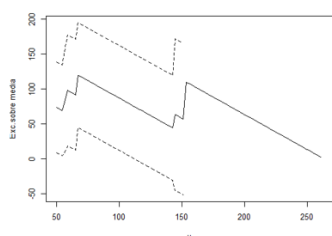
- μ es el parámetro que se corresponde con la localización → 37.7934
- σ es el parámetro que se corresponde con la escala → 28.9358
- ξ es el que se corresponde con la forma o índice de cola, es decir, indica el tamaño de la cola de la distribución → 0.6384
- Valor mínimo → 52.233
- Número de iteraciones → 37

MODELO Poisson-Generalizada de Pareto

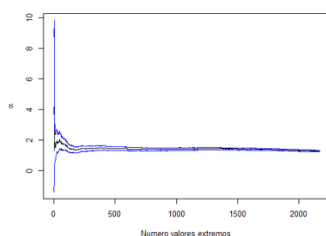
- μ es el parámetro que se corresponde con la localización → 178.82
- ξ es el que se corresponde con la forma o índice de cola, es decir, indica el tamaño de la cola de la distribución → -0.6398
- Valor mínimo → 38.826
- Número de iteraciones → 19

Para el modelo de Poisson-Generalizada de Pareto tenemos una cola corta puesto que obtenemos un $\xi < 0$. Adicionalmente a lo anterior realizamos la estimación basándonos en otros índices de cola

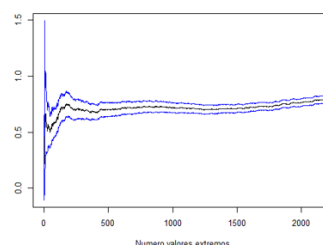
Exceso sobre la media



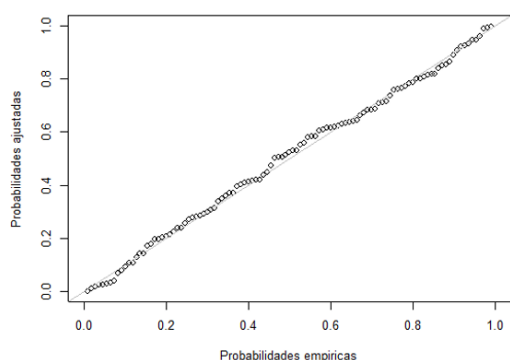
Intervalo de confianza de $\alpha=1/\text{índice de cola}$



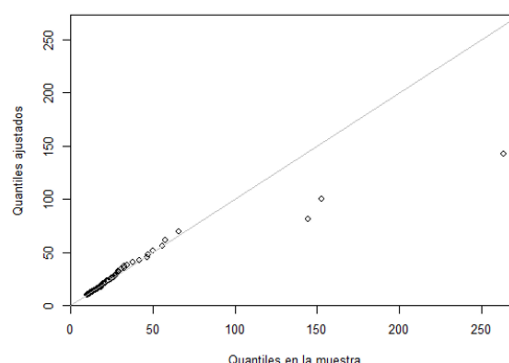
Intervalo de confianza asintótico del índice de cola



A fin de poder validar nuestros modelos procedemos a su validación en donde validaremos los cuartiles de la muestra, así como relacionaremos las probabilidades empíricas con las teóricas para determinar si el impacto producido por las pérdidas teóricas es capaz de explicar las pérdidas incurridas.



- Se observa como existe una buena relación entre las probabilidades ajustadas con las realmente obtenidas tras nuestro análisis

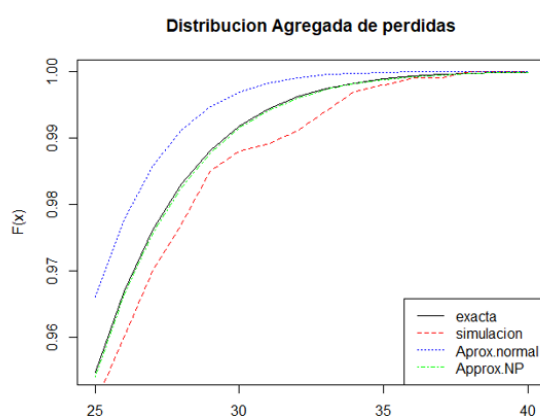


- El ajuste de los quantiles muestrales están ajustados con los obtenidos en el total de la población. Están concentrados en torno a la izquierda

5. Comunicación de resultados

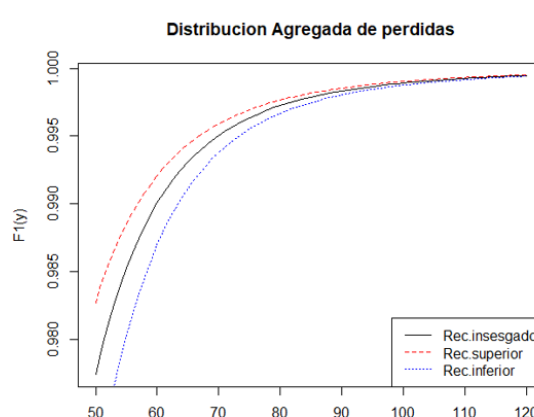
Todo lo expuesto anteriormente nos permite analizar el impacto producido por las pérdidas agregadas. Para ello basaremos nuestros análisis por medio de dos vías. Por una parte, a través del modelo Gamma y en donde veremos que la mejor aproximación hacia la Distribución de Pérdidas Agregadas la ofrece la Distribución Normal-Power y, por otra parte, a través del método Pareto en donde calcularemos la función de distribución discretizándola.

CASO 1: GAMMA



- N sigue una distribución de Poisson y X una Gamma
- Establecemos los momentos de orden 1 Gamma para la media y orden 2 para la varianza, así como el Coeficiente de Asimetría
- Realizamos el proyecto simulación, la aproximación a la Normal-Power, la Normal y la exacta

CASO 2: PARETO



- N sigue una distribución de Pareto (10) y sigue una Gamma (3,2)
- Calculamos la función de distribución discretizándola
- Realizamos el análisis a través del método insesgado, superior e inferior

6. Apéndice: Medición del riesgo extremo: Value At Risk (VaR)

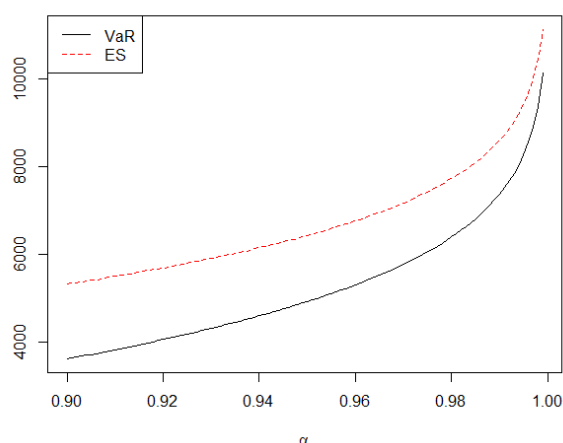
El concepto de VaR se considera la medida estándar en la medición y la gestión del riesgo de mercado, mide las pérdidas máximas estimadas en el valor de mercado de una cartera, es decir, calcula una eventual pérdida extrema para un determinado período de tiempo utilizando una medida de volatilidad de los últimos 'n' días.

Ejercicio 1:

El VaR al 95% es de 4921.473
El VaR al 99% es de 7374.852

La ES al 95% es de 6425.766
La ES al 99% es de 8594.771

La representación gráfica de los resultados obtenidos sería el siguiente



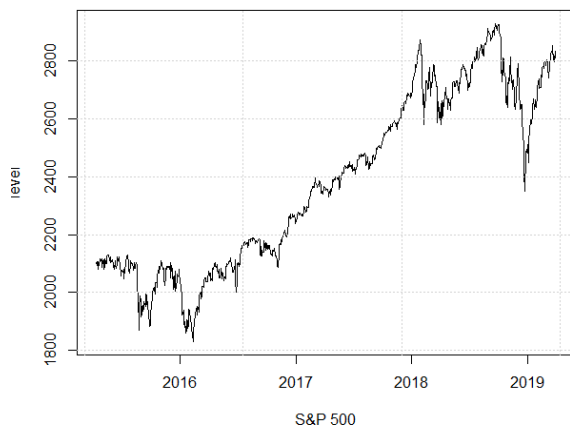
Ejercicio 2:

El VaR con la t-student al 95% es de 32897.07
El VaR con la t-student al 99% es de 46526.96

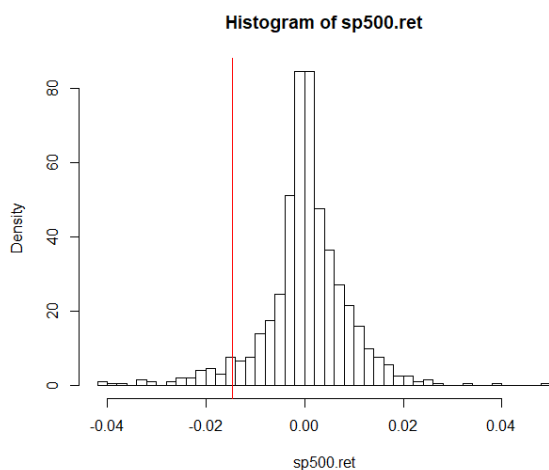
La ES con la t-student al 95% es de 5705.099
La ES con la t-student al 99% es de 9329.63

Ejercicio 3:

La representación gráfica resultante de seleccionar 1000 registros sobre el índice S&P 500

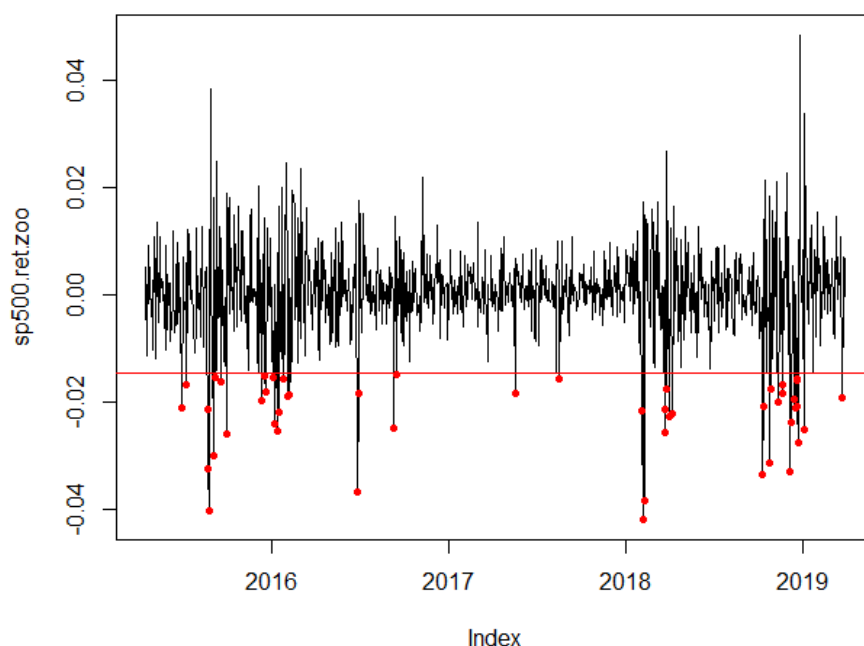


La representación del histograma para ese período particular utilizando 50 rupturas vendría de la siguiente manera



El VaR empírico derivado de la simulación histórica al 95% es de 292.3274. A su vez el ES empírico derivado de la simulación histórica al 95% es de 451.9102.

Los intervalos de confianza al 95% para las estimaciones obtenidas utilizando 10.000 repeticiones mediante Bootstrap son (250.5, 348.6) y (411.4, 494.1). La representación gráfica vendría de la siguiente manera



7. CONCLUSIONES

A través de esta práctica he podido realizar análisis exhaustivos acerca del riesgo operativo sobre una compañía con la distribución de datos de la empresa 'Vikings_Fish' los cuales me han permitido no solamente comparar los resultados teóricos con los empíricos sino también realizar un análisis sobre el impacto derivado de las pérdidas agregadas en base a diferentes metodologías y analizar el Value at Risk (VaR) de la empresa.

8. OPINION PERSONAL

El riesgo operacional es absolutamente imprescindible para las entidades financieras. La medición del riesgo operacional con la Circular del Banco de España 3/2008, Sección Primera establece dos tipos de medición, que de menor a mayor grado de sofisticación y sensibilidad al riesgo y de mayor a menor nivel de requerimientos de recursos propios son el resultado de aplicar el método básico (BIA, Basic Indicator Approach) o el método avanzado (AMA, Advanced Measurement Approach).

Según mi experiencia en el Banco Popular Español y en la Caja Rural de Navarra he de decir que el avance y el desarrollo de metodologías como el AMA es el principal reto al que se encuentran las entidades financieras en cuanto a la medición de riesgo operacional. La evolución histórica del tratamiento de riesgo operacional y el enfoque de los reguladores nos demuestra que un

reconocimiento adecuado de eventos de pérdidas es el principal punto de partida para un adecuado tratamiento y medición del riesgo. Considero que no solamente debemos ceñirnos en trabajar con bases de datos internas sino también considerar fuentes externas que supongan complementar las pérdidas extremas que a nivel metodológico se sitúan en las colas de la distribución de severidad.

9. BIBLIOGRAFÍA

Sonia de Paz. CUNEF, 2019. Apuntes tomados en clase y diapositivas sobre riesgo operacional y valores extremos.

Banco de España [2008]: “Circular 3/2008 a entidades de crédito, sobre determinación y control de los recursos propios mínimos”.

Hoffman, D.G. [1998]: “New Trends in Operational Risk Measurement

The Treatment of Expected Losses by Banks Using the AMA under the Basel II Framework, N°7, Basilea, noviembre, 2005