

# TRABAJO FIN DE MÁSTER

---

**Título: Cuantificación del riesgo de pérdida: Cópulas y Probit  
Multivariante**

**Autor: Carla de Pedro Cascón**

**Tutor/res: Catalina Bolancé Losilla**

**Curso: 2016/2017**

Facultad de Economía y Empresa

Universidad de Barcelona

Trabajo Fin de Máster

Máster en Ciencias Actuariales y Financieras

**CUANTIFICACIÓN DEL  
RIESGO DE PÉRDIDA:  
CÓPULAS Y PROBIT  
MULTIVARIANTE**

Autor: Carla de Pedro Cascón

Tutor/es: Catalina Bolancé Losilla

El contenido de este documento es de exclusiva responsabilidad del autor, quien declara que no ha incurrido en plagio y que la totalidad de referencias a otros autores han sido expresadas en el texto.



# Cuantificación del riesgo de pérdida: Cópulas y Modelo Probit Multivariante

Carla de Pedro Cascón

2 de julio de 2017

## Resumen

El presente trabajo de final de máster se centra en modelizar los costes de una cartera de seguros no vida, teniendo en cuenta la probabilidad de que se produzca el siniestro. Una aseguradora trata con miles de contratos en los que se compromete a cubrir determinados accidentes. Por tanto, será esencial que esté preparada para cumplir sus obligaciones. En este sentido, para las compañías de seguros la cuantificación del riesgo de pérdida de una forma eficiente será clave para cumplir con los requerimientos de capital de solvencia.

**Palabras clave:** Modelo Probit Multivariante, Cópulas, Seguros no vida, *Value at risk* VaR, Pérdida extrema, Dependencia en las colas.

## Abstract

The aim of this master thesis is focus on modeling the cost of a nonlife portfolio, considering the probability that a claim occurs. An insurance entity deals with thousands of contracts and compromises to cover the amount of claims. Thus, it is essential to be prepared to accomplish the liabilities. So, for the insurance companies, the efficient quantification of the risk of loss is key to cope with the design imposed by solvency capital requirements.

**Keywords:** Probit Multivariate Model, Copulas, Nonlife insurance policies, *Value at risk*, Extrem loss, Tail dependence.



# Índice

<b>1. Introducción</b>	<b>9</b>
<b>2. Modelos Lineales Generalizados</b>	<b>11</b>
2.1. Estructura de los GLM . . . . .	12
2.2. Estimación de los parámetros en los modelos GLM . . . . .	14
2.3. Modelos Binomiales - Modelos para variables binarias . . . . .	14
2.4. Modelo Probit Multivariante . . . . .	16
<b>3. Cópulas</b>	<b>18</b>
3.1. Tipos de Cópulas . . . . .	19
3.1.1. Cópulas elípticas . . . . .	19
3.1.2. Cópula arquimediana . . . . .	21
3.2. Medidas de dependencia . . . . .	24
3.2.1. Coeficiente lineal de Pearson: . . . . .	24
3.2.2. Medidas de dependencia utilizando cópulas . . . . .	25
3.3. Estimación de los Parámetros y Criterios de Selección de la Cópula . . . . .	26
3.3.1. Estimación de los parámetros . . . . .	26
3.3.2. Criterios de Selección . . . . .	28
3.4. Medidas de riesgo . . . . .	28
3.5. Estimación del riesgo integral del asegurado . . . . .	30
<b>4. Contexto práctico</b>	<b>32</b>
4.1. Descripción de los Datos . . . . .	32
4.2. Aplicación del Probit Multivariante . . . . .	35
4.3. Perfiles . . . . .	37
4.4. Aplicación de las cópulas . . . . .	38
4.4.1. Análisis del caso trivariante: Daños materiales - Daños corporales - Hogar	39
4.4.2. Caso bivariante: Daños corporales - Hogar . . . . .	44
4.4.3. Caso bivariante: Daños materiales - Daños corporales . . . . .	48
4.4.4. Caso bivariante: Daños materiales - Hogar . . . . .	51
4.4.5. Caso univariante: Daños materiales . . . . .	54
4.4.6. Caso univariante: Daños corporales . . . . .	55
4.4.7. Caso univariante: Hogar . . . . .	57
4.5. Riesgo de pérdida integrado . . . . .	59
<b>5. Conclusiones</b>	<b>62</b>

Anexo	63
Referencias	110

## Índice de figuras

1. Comparación Logit vs Probit	16
2. Relación entre leyes marginales, cópula y leyes conjuntas	18
3. Función de distribución, densidad y líneas de nivel de la cópula Gaussiana, $\rho=0.5$	23
4. Función de distribución, densidad y líneas de nivel de la cópula t-Student, $\rho=0.5$ , $\nu = 4$	23
5. Función de distribución, densidad y líneas de nivel de la cópula Gúmbel, $\theta=4$	24
6. Gráfico VaR-TVaR	29
7. Histogramas: Daños materiales, Daños corporales, Hogar	40
8. Histogramas: Daños corporales, Hogar	45
9. Histogramas: Daños materiales, Daños corporales	48
10. Histogramas: Daños materiales, Hogar	51
11. Histograma Daños Materiales	54
12. Histograma Daños corporales	56
13. Histograma Hogar	58

## Índice de tablas

1. Funciones enlace. <i>Fuente: Elaboración propia</i>	13
2. Clasificación cópulas	22
3. Explicación de las variables	32
4. Estadísticos descriptivos para las variables dependientes - Número de siniestros	33
5. Estadísticos descriptivos para las variables dependientes - Costes	34
6. Estadísticos descriptivos para las variables explicativas	34
7. Matriz correlación entre los diferentes riesgos	35
8. Estimación de los parámetros para cada línea de negocio	36
9. Perfiles	37
10. Resultados del cálculo de probabilidad	38
11. Estadísticos descriptivos caso trivariante: Daños materiales - Daños corporales - Hogar	39
12. Estimación parámetros caso trivariante: Daños materiales - Daños corporales - Hogar	41



13.	Elección cópula caso trivariante: Daños materiales - Daños corporales - Hogar .	41
14.	Coefficiente de dependencia de la cola caso trivariante: Daños materiales - Daños corporales - Hogar . . . . .	42
15.	Elección marginales caso trivariante: Daños materiales - Daños corporales - Hogar	43
16.	VaR- Cópula t-Student, caso trivariante: Daños materiales - Daños corporales - Hogar . . . . .	44
17.	Estadísticos descriptivos caso bivariante: Daños corporales - Hogar . . . . .	45
18.	Estimación parámetros caso bivariante: Daños corporales - Hogar . . . . .	45
19.	Estimación cópula caso bivariante: Daños corporales - Hogar . . . . .	46
20.	Coefficiente de dependencia de la cola caso bivariante: Daños corporales - Hogar	46
21.	Elección marginales caso bivariante: Daños corporales - Hogar . . . . .	47
22.	VaR- Cópula Gaussiana, caso bivariante: Daños corporales - Hogar . . . . .	47
23.	Estadísticos descriptivos caso bivariante: Daños materiales - Daños corporales .	48
24.	Estimación parámetros caso bivariante: Daños materiales - Daños corporales . .	49
25.	Elección cópula caso bivariante: Daños materiales - Daños corporales . . . . .	49
26.	Coefficiente de dependencia de la cola caso bivariante: Daños materiales - Daños corporales . . . . .	49
27.	Elección marginales caso bivariante: Daños materiales - Daños corporales . . . .	50
28.	VaR- Cópula t-Student, caso bivariante: Daños materiales - Daños corporales . .	50
29.	Estadísticos descriptivos caso bivariante: Daños materiales - Hogar . . . . .	51
30.	Estimación parámetros caso bivariante: Daños materiales - Hogar . . . . .	52
31.	Elección cópula caso bivariante: Daños materiales - Hogar . . . . .	52
32.	Coefficiente de dependencia de la cola caso bivariante: Daños materiales - Hogar	52
33.	Elección marginales caso bivariante: Daños materiales - Hogar . . . . .	53
34.	VaR- Cópula Gumbel, caso bivariante: Daños materiales - Hogar . . . . .	53
35.	Estadísticos descriptivos caso univariante: Daños materiales . . . . .	54
36.	Elección marginales caso univariante: Daños materiales . . . . .	55
37.	VaR, caso univariante: Daños materiales . . . . .	55
38.	Estadísticos descriptivos caso univariante: Daños corporales . . . . .	56
39.	Elección marginales caso univariante: Daños corporales . . . . .	56
40.	VaR, caso univariante: Daños corporales . . . . .	57
41.	Estadísticos descriptivos caso univariante: Hogar . . . . .	57
42.	Elección marginales caso univariante: Hogar . . . . .	58
43.	VaR, caso univariante: Hogar . . . . .	59
44.	Resultados riesgo de pérdida con $\alpha=0.995$ integrado para los perfiles . . . . .	60
45.	TVaR- Cópula t-Student, caso trivariante: Daños materiales - Daños corporales - Hogar . . . . .	63
46.	TVaR- Cópula Gaussiana, caso bivariante: Daños corporales - Hogar . . . . .	63

47.	TVaR- Cópula t-Student, caso bivalente: Daños materiales - Daños corporales .	63
48.	TVaR- Cópula Gumbel, caso bivalente: Daños materiales - Hogar . . . . .	64
49.	TVaR, caso univalente: Daños materiales . . . . .	64
50.	TVaR, caso univalente: Daños corporales . . . . .	64
51.	TVaR, caso univalente: Hogar . . . . .	65

# 1. Introducción

Uno de los principales retos en el ámbito financiero-actuarial, es adoptar métodos que permitan medir el riesgo y su incidencia en la viabilidad de la empresa. Las instituciones financieras y aseguradoras deben tener muy clara su estructura organizacional en lo que respecta a la gestión de riesgos. Tanto Basilea III, para el marco financiero, como Solvencia II, para el marco asegurador, exigen unos niveles de exposición al riesgo en los que puede incurrir la entidad. De ahí la importancia de garantizar el riesgo asumido.

En este trabajo, se analiza la distribución asociada con la pérdida total mediante un vector aleatorio multivariante  $(L_1, L_2, \dots, L_d)$  de pérdidas dependientes o factores de riesgo. Definiremos la pérdida total como  $L = L_1 + L_2 + \dots + L_d$  y el objetivo será estimar el riesgo de pérdida de una cartera de seguros no vida. La distribución de las pérdidas en una cartera de seguros no vida tiene una fuerte asimetría hacia la derecha, puesto que la mayor parte de los asegurados no tienen siniestros, o en su caso, son de un importe bajo. Nuestro interés será considerar aquellos que suponen una pérdida elevada, para ser capaz de cubrirlos. Dado que, en el entorno de los seguros se puede incurrir en pérdidas de mucho valor, es fundamental una buena cuantificación del riesgo, de manera que cuanto más precisa sea su medición, más garantía de estabilidad económica obtendremos. Así pues, Solvencia II exige que la entidad esté en situación de cumplir con sus obligaciones frente a los tomadores y beneficiarios con una probabilidad de un 99,5 % como mínimo.

La pérdida global de la cartera suele estar relacionada con la dependencia entre los diferentes tipos de riesgo de los que está compuesta. Una manera sencilla de medir la dependencia es utilizar la correlación lineal, pero se basa en supuestos de normalidad, lo que puede implicar que se infravalore el nivel de riesgo de los eventos extremos que están asociados a las pérdidas más elevadas. Luego, si las pérdidas agregadas son asimétricas positivas, como en las carteras de seguros no vida, estas medidas de dependencia pueden no relacionar los riesgos de la forma más eficiente.

En este sentido, determinar la estructura de dependencia entre los diferentes riesgos que afectan a la cartera por medio de una forma alternativa, como son las cópulas, puede ayudar a mejorar la estimación de la distribución de las pérdidas esperadas para determinar el nivel de provisiones necesario y así poder hacer frente a unas posibles pérdidas.

En este trabajo veremos cómo seleccionar la cópula y cómo estimar las distribuciones marginales cuando tenemos valores extremos. Calcularemos el riesgo de pérdida de una cartera usando cópulas paramétricas, condicionado a que se declare el siniestro. Esta condición se de-

terminará a partir del Modelo Probit Multivariante, un tipo de Modelo Lineal Generalizado que nos informará de la probabilidad que un asegurado, con unas determinadas características, declare el siniestro. Para estos cálculos se han utilizado los programas estadísticos *SAS* y *R Commander*.

La muestra elegida presenta dos tipos de seguro:

1. El seguro de hogar, que tiene por objeto cubrir los daños que ocurran en una vivienda ya sean de continente (todo lo relacionado con el inmueble y la construcción) como del contenido (todo lo que se tiene dentro del inmueble).
2. El seguro de automóvil, que cubrirá accidentes propios o causados por un tercero. Para este tipo de seguro podemos dividir el riesgo al que está expuesto el asegurado en:
  - a) Los daños corporales que son los producidos a las personas, como lesiones, días de baja médica y hospitalización, rehabilitación, secuelas. . .
  - b) Los daños materiales relacionados a las cosas materiales y a los animales.

Por lo tanto, trabajaremos con una cartera que está compuesta de tres líneas de negocio, dado que se considera interesante separar los daños corporales de los daños materiales. Entonces, calcularemos la pérdida integrada como sigue:

$y_{md}=1$ , si declara siniestro de daños materiales	$C_{md}$ = costes daños materiales
$y_{bi}=1$ , si declara siniestro de daños corporales	$C_{bi}$ = costes daños corporales
$y_{home}=1$ , si declara siniestro de hogar	$C_{home}$ = costes siniestros de hogar

$$\begin{aligned}
 L = & \rho(C_{md}, C_{bi}, C_{home}) \times Pr(y_{md} = 1, y_{bi} = 1, y_{home} = 1) + \\
 & \rho(C_{md}, C_{bi}) \times Pr(y_{md} = 1, y_{bi} = 1, y_{home} = 0) + \\
 & \rho(C_{md}, C_{home}) \times Pr(y_{md} = 1, y_{bi} = 0, y_{home} = 1) + \\
 & \rho(C_{bi}, C_{home}) \times Pr(y_{md} = 0, y_{bi} = 1, y_{home} = 1) + \\
 & \rho(C_{md}) \times Pr(y_{md} = 1, y_{bi} = 0, y_{home} = 0) + \\
 & \rho(C_{bi}) \times Pr(y_{md} = 0, y_{bi} = 1, y_{home} = 0) + \\
 & \rho(C_{home}) \times Pr(y_{md} = 0, y_{bi} = 0, y_{home} = 1)
 \end{aligned}$$

donde  $\rho$  es la medida de riesgo univariante que depende de los costes producidos.

## 2. Modelos Lineales Generalizados

Los modelos lineales generalizados, también conocidos por el acrónimo GLM (*Generalized linear models*), por sus siglas en inglés, representan una importante parte de los modelos de regresión. Gracias a sus ventajas, es una de las técnicas predictivas más utilizadas en el campo actuarial, sobre todo en la modelización para las compañías de seguros no vida.

Son utilizados para analizar la relación entre la variable respuesta, sea esta la frecuencia o la severidad de los siniestros, y las variables explicativas. El caso más simple es el modelo de regresión lineal clásico, que consta de tres supuestos principales:

- Los errores se distribuyen según una Normal.
- Homocedasticidad<sup>1</sup>.
- La relación entre la variable dependiente y la(s) variable(s) explicativa(s) es lineal.

El modelo de regresión lineal múltiple será:

$$y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + \varepsilon_i, \quad i = 1, \dots, n, \quad (2.1)$$

donde los residuos se comportan como<sup>2</sup>,

$$\varepsilon_i \sim \mathcal{N}(0, \sigma^2).$$

El comportamiento de la variable  $y_i$  viene explicado por un modelo o relación en el que se distingue una parte determinista (integrada por las variables explicativas) y una parte aleatoria (término error).

Si tomamos la esperanza de la ecuación (2.1):

$$E(y_i) = \mu = \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki}. \quad (2.2)$$

Se observa que la relación entre la variable dependiente y la combinación lineal de las variables explicativas es la función identidad. Sin embargo, bajo las restricciones del modelo clásico, nos encontramos en muchas ocasiones que uno o varios supuestos no se cumplen. Por ello, se utilizan los GLM, siendo una extensión flexible del modelo lineal ordinario en dos direcciones:

---

<sup>1</sup>La varianza es constante

<sup>2</sup>Los residuos que así se distribuyen se conocen como ruido blanco

- La distribución de la variable respuesta se elige de la familia exponencial. Entonces, la distribución de la variable respuesta no necesita ser normal.
- La transformación de la esperanza de la variable respuesta está relacionada con la combinación lineal de las variables explicativas.

Una consecuencia directa de permitir que la distribución de la variable respuesta forme parte de la familia exponencial, es que permite que ésta sea heterocedástica<sup>3</sup>, de hecho es lo que suele ocurrir.

Es importante destacar que los datos relacionados con la frecuencia y la severidad de los siniestros no suelen seguir una distribución normal, por consiguiente, los modelos GLM son de mucha utilidad.

## 2.1. Estructura de los GLM

Los Modelos Lineales Generalizados tienen tres componentes básicos:

1. **Componente aleatorio:** Dada una variable respuesta  $y$ , su distribución se caracteriza por:

$$f(y) = \exp \left\{ \frac{y\theta - b(\theta)}{\phi/\omega} + c(y, \phi) \right\}; \quad g(\mu) = x'\beta. \quad (2.3)$$

- La función de densidad  $f(y)$  especifica que la distribución de la variable respuesta  $y$  es de la familia exponencial.
- $\phi$  es un parámetro de dispersión.
- La elección de  $b(\theta)$  determina la distribución de la variable respuesta. La función  $c(\cdot)$  también es conocida.
- $\omega$  es una ponderación<sup>4</sup>.
- $\theta$  es el parámetro canónico.
- La ecuación  $g(\mu)$ , denominada función enlace, establece cómo varía la media respecto las variables explicativas  $x$ .

---

<sup>3</sup>La varianza variará con la media

<sup>4</sup> $\omega=1$  si son datos individuales,  $\omega=n_i$  si son datos promedios,  $\omega=1/n$  si son datos como suma de  $n_i$  respuestas individuales

En el modelo lineal, la relación entre la media y las variables explicativas es  $\mu = x'\beta$ . Mientras que en los GLM, se generaliza a  $g(\mu) = x'\beta$ . Donde  $g$  es una función monótonica y diferenciable.

2. **Componente sistemático:** está asociado a las variables explicativas utilizadas en la función predictora lineal. Por lo tanto, el valor del predictor lineal es:

$$\eta_i = \sum_{j=1}^k x_{ij}\beta_j = \beta_0 + \beta_1x_{i1} + \beta_2x_{i2} + \cdots + \beta_kx_{ik}, \quad \forall_i = 1, \dots, n. \quad (2.4)$$

- Cada  $\beta_j$  está asociada a una variable  $j$  e individuo  $i$ .
- $x_{ij}$  será el valor de la variable  $j$  en el individuo  $i$ .

3. **Componente ligadura:** Se trata de la función *link* o enlace y relaciona el valor esperado de la variable dependiente con el componente sistemático. Es decir, especifica una función  $g(\cdot)$  que relaciona la  $E(y_i)$  con el predictor lineal:

$$\eta_i = g(\mu_i) = \sum_{j=1}^k \beta_j x_{ij}. \quad (2.5)$$

La función  $g(\mu)$  es una función conocida, monótona y diferenciable de  $\eta$ , entonces:

$$\mu_i = g^{-1}(\eta_i). \quad (2.6)$$

En la siguiente tabla podemos ver algunos de los modelos más utilizados con sus respectivas ligaduras:

Modelo	Ligadura	$\eta_i = g(\mu_i)$	$\mu_i = g^{-1}(\eta_i)$	Dominio de $Y_i$	$Var(\mathbf{Y}_i   \eta_i)$
Normal	Identidad	$\mu_i$	$\eta_i$	$(-\infty, +\infty)$	$\phi$
Binomial	Logit	$\ln \frac{\mu_i}{1-\mu_i}$	$\frac{\exp(\eta_i)}{1+\exp(\eta)}$	$\frac{0,1,\dots,n_i}{n_i}$	$\frac{\mu_i(1-\mu_i)}{n_i}$
Binomial	Probit	$\Phi^{-1}(\mu_i)$	$\Phi(\eta_i)$	$\frac{0,1,\dots,n_i}{n_i}$	$\frac{\mu_i(1-\mu_i)}{n_i}$
Poisson	Log	$\ln(\mu_i)$	$\exp(\eta_i)$	$0,1,2,\dots$	$\mu_i$
Gamma	Inversa	$\mu_i^{-1}$	$\eta_i^{-1}$	$(0, \infty)$	$\phi\mu_i^2$

Tabla 1: Funciones enlace. *Fuente: Elaboración propia*

## 2.2. Estimación de los parámetros en los modelos GLM

Cuando la variable dependiente se distribuye según una distribución de la familia exponencial, los parámetros suelen ser estimados por máxima verosimilitud (MV). La idea es encontrar aquellos valores de los parámetros  $\beta_j$  que hacen máxima la probabilidad de que los datos tratados queden caracterizados por dichos parámetros.

Para maximizar la función de verosimilitud el método más usual es aplicar derivadas parciales respecto los parámetros de interés y encontrar las raíces de los resultados de las ecuaciones. Antes es necesario linealizar la función, lo que bastará con tomar logaritmos. Puesto que resolver esta ecuación es complicado, se aplican métodos de aproximación por procedimientos iterativos como pueden ser el algoritmo de *Newton-Raphson* o el método de *Scoring Fisher*.

La base para explicar los GLM se ha obtenido de los libros de De Jong y Heller (2008b), Frees (2010b) y McCullagh y Nelder (1983b) y del artículo de Cayuela (2010a).

## 2.3. Modelos Binomiales - Modelos para variables binarias

En general los modelos discretos tienen el objetivo de explicar y predecir el comportamiento de variables, que en ciertas ocasiones son discretas, es decir, hay un número limitado de posibilidades o alternativas (Madala, 1983).

En los modelos de respuesta binaria, la variable respuesta  $y_i$  puede tomar dos valores, codificándolos usualmente como 1 para la categoría de interés y 0 para la otra. De este modo, para cada individuo  $i$  se puede definir una variable  $y_i$  para que tome dos posibles valores:

$$\begin{aligned}Pr(y_i = 1) &= F(x'_i, \beta) \\Pr(y_i = 0) &= 1 - F(x'_i, \beta),\end{aligned}$$

donde  $x'_i = (1, x_{i1}, x_{i2}, \dots, x_{ik})'$  es un vector de una serie de factores que explican la variable dependiente y  $\beta = (\beta_0, \beta_1, \beta_2, \dots, \beta_k)'$  es el vector de parámetros que refleja el impacto que  $x_i$  tiene sobre la probabilidad.

De manera que  $y_i \sim \text{Binomial}(1, \pi)$  o de forma equivalente  $y_i \sim \text{Bernoulli}(\pi)$ .



La esperanza del valor de la variable dicotómica  $y_i \in \{0, 1\}$  será:

$$E[y_i] = Pr(y_i = 1) = \pi_i. \quad (2.7)$$

Este modelo supone que la probabilidad aumenta linealmente con  $x_i$ . Un modo de estimar el modelo con variable dependiente binaria es a partir del modelo de probabilidad lineal, que equivale al que se definió en (2.1), pero con la variable dependiente binaria. Esta forma de estimar se usa debido a la fácil interpretación de los parámetros, pero tienen varios inconvenientes:

- El modelo en ocasiones será incorrecto puesto que la probabilidad por definición está acotada entre 0 y 1, pero las ecuaciones lineales no tienen límites, varían entre  $+\infty, -\infty$ . Por lo tanto, es posible que el modelo lineal de probabilidad estime probabilidades mayores de 1 o menores de 0.
- Los modelos lineales asumen homocedasticidad, pero la varianza de la variable respuesta depende de la media y ésta varía con las observaciones. Por lo tanto, no se tiene en cuenta la heterocedasticidad existente.
- Falta de normalidad en la variable  $y_i$  y por consiguiente también en los errores, ya que tanto  $y_i$  como  $\varepsilon_i$  se distribuyen según una Bernoulli.
- La relación entre  $\pi(x)$  y  $x$  debería ser lineal. Pero en muchas coyunturas no es así, puesto que el cambio en  $x$  tiene menor impacto cuando  $\pi$  es más cercano a 0 o 1 que cuando se encuentra en el centro.

Para combatir los inconvenientes de los modelos lineales de probabilidad, consideraremos unos modelos alternativos, de forma que relacionarán la variable endógena  $y_i$  con el vector de las variables explicativas  $x_i$  a través de una función de distribución. Entonces se expresa la esperanza de la variable respuesta como una función de las variables explicativas,  $\pi_i = \pi(x'_i\beta) = Pr(y_i = 1)$ . Nos centraremos en dos casos especiales de la función  $\pi(\cdot)$ :

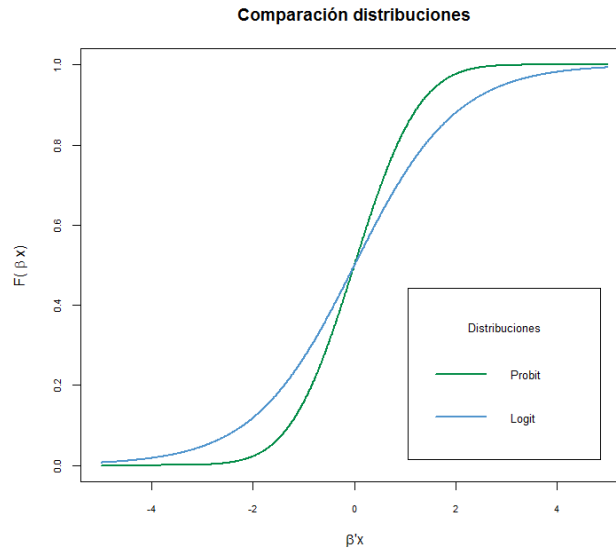
- **Logit:** supone que  $\pi(\cdot)$  es la función de distribución logística:

$$\pi(x'_i\beta) = \frac{1}{1 + e^{-x'_i\beta}} = \frac{e^{x'_i\beta}}{1 + e^{x'_i\beta}}. \quad (2.8)$$

- **Probit:** supone que  $\pi(\cdot)$  es una función de distribución de la normal estándar:

$$\pi(x'_i\beta) = \Phi(x'_i\beta) = \int_{-\infty}^{x'_i\beta} \phi(t)dt = \int_{-\infty}^{x'_i\beta} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt. \quad (2.9)$$

Figura 1: Comparación Logit vs Probit



*Fuente: Elaboración propia*

En este caso,  $\pi$  es no lineal para ambas distribuciones (Logit y Probit). Estas dos funciones son muy similares, aunque como podemos observar en la Figura 1, la función logística crece de forma más moderada. Además, la probabilidad de la variable respuesta estará entre 0 ó 1, dado que el uso de una función de distribución garantiza que el resultado esté acotado entre estos valores.

La base para redactar los Modelos Binomiales se ha obtenido de los libros de Tse (2009b), Greene (2003b), McCullagh y Nelder (1983b), y de los artículos de Flórez y Rincón (2012a) y Schmidheiny (2016a).

## 2.4. Modelo Probit Multivariante

El Modelo Probit Multivariante es una extensión natural del Modelo Probit Univariante. Se utiliza en los casos que hay más de una variable respuesta binaria, de forma que considera más de una ecuación y permite que los errores estén correlacionados.

Para este modelo, asumimos que cada sujeto  $i$  tiene un vector de covariadas y además

está relacionado con  $d$  respuestas o categorías dicotómicas diferentes. Entonces, denominamos  $M=(y_{i1}, \dots, y_{id})'$  al conjunto de respuestas dicotómicas 0/1 observadas para las  $d$  variables del  $i$ -ésimo sujeto,  $i=1, \dots, n$ , para determinar la posibles combinaciones de las  $d$  diferentes categorías, será  $2^d$ :

$$\begin{bmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{id} \end{bmatrix} = \begin{bmatrix} x'_{i1}\beta_1 \\ x'_{i2}\beta_2 \\ \vdots \\ x'_{id}\beta_d \end{bmatrix} + \begin{bmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \\ \vdots \\ \varepsilon_{id} \end{bmatrix}, \quad (2.10)$$

donde  $x_{ij}=(1, x_1, x_2, \dots, x_k)'$  es el vector de covariadas, siendo el mismo para todas las categorías;  $\beta_j=(\beta_{j0}, \beta_{j1}, \beta_{j2}, \dots, \beta_{jk})'$  es el correspondiente vector de parámetros desconocidos que son los que debemos estimar y  $\varepsilon_{id}$  es el componente estocástico, que se considera como el conjunto de factores inobservables.

Cada  $\varepsilon_{id}$  está compuesto de una  $d$ -variada distribución Normal con media cero y varianza estandarizada a la unidad, donde  $\varepsilon \sim \mathcal{N}(0, \Omega)$ ,  $\Omega$  es la matriz de covarianzas simétrica y con diagonal igual a 1:

$$\Omega = \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1d} \\ \rho_{21} & 1 & \cdots & \rho_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{d1} & \rho_{d2} & \cdots & 1 \end{bmatrix},$$

siendo  $\rho_{id}$  la correlación entre el  $i$ -ésimo componente estocástico y los  $d$  tipos de variables respuesta. Además sabemos que la matriz es simétrica, por lo tanto  $\rho_{id}=\rho_{di}$ .

La notación del Modelo Probit Multivariante, deriva directamente del Modelo Probit Univariante. Entonces para el conjunto de las posibles  $d$  categorías:

$$Pr(y_1 = 1, y_2 = 1, \dots, y_d = 1) = \Phi_d(x'_i\beta_1, x'_i\beta_2, \dots, x'_i\beta_d; \Omega), \quad (2.11)$$

donde  $\Phi(\cdot)$  es la función de distribución Normal estándar y  $\Omega$  es la matriz de covarianza.

Para la elaboración de este apartado, se han utilizado los manuales de Greene (2003b) y aa1985 y los artículos de Young *et al.* (2006a) y Chib y Greenberg (1998a).

### 3. Cópulas

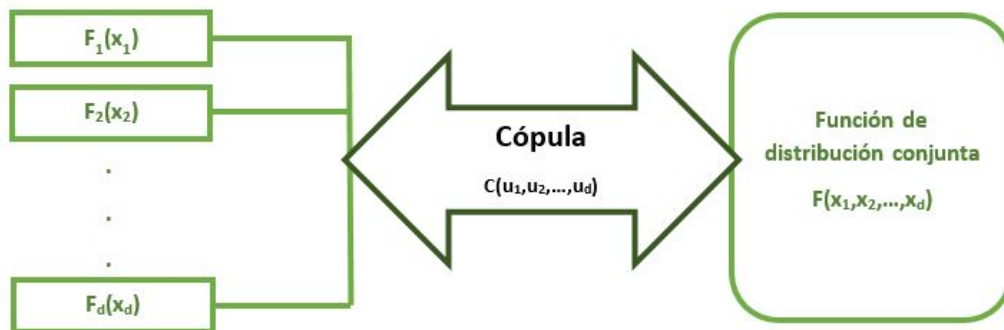
El concepto cópula proviene del latín y significa “unión”. Proporciona una forma de determinar la estructura de dependencia, siendo por tanto una herramienta muy útil para modelar un conjunto de factores.

En el contexto financiero-actuarial, la pérdida total, a la que denominamos  $L$ , suele estar generada por sub-pérdidas que provienen de diferentes  $d$  riesgos. De forma que,  $L = L_1 + L_2 + \dots + L_d$  y consecuentemente, la pérdida total depende de la relación que exista entre los  $d$  riesgos. Una particularidad que hace muy interesante a las cópulas es que miden tanto la dependencia lineal como la no lineal.

**Definición 3.1 (cópula)** Una cópula  $d$ -dimensional es una función de distribución en  $[0, 1]^d$  con marginales uniformes estándar. Siendo  $C(u) = C[u_1, u_2, \dots, u_d]$ . Por lo tanto,  $C : [0, 1]^d \rightarrow [0, 1]$ .

Entonces, una cópula es una función de distribución multivariante definida sobre el dominio  $[0, 1]^d$ , tal que cada distribución marginal es uniforme en el intervalo  $[0, 1]$ . En otras palabras, será una función que une una distribución de probabilidad multivariante con una serie de distribuciones de probabilidad marginales univariante y de este modo determina la dependencia de la distribución conjunta con las distribuciones marginales.

Figura 2: Relación entre leyes marginales, cópula y leyes conjuntas



Fuente: Elaboración propia

El teorema de Sklar está implícito en las aplicaciones de las cópulas. En primer lugar, muestra que todas las funciones de distribución multivariadas contienen cópulas. Y en segundo lugar, las cópulas se pueden emplear para construir distribuciones multivariadas a partir de las distribuciones marginales, puesto que cada función de distribución se puede descomponer en sus marginales y como mínimo en una cópula<sup>5</sup>.

**Teorema 3.1 (Sklar, 1959)** *Sea  $F(x)$  una función de distribución conjunta, con marginales  $F_1(x_1), F_2(x_2), \dots, F_d(x_d)$ . Entonces existe una cópula  $C$  tal que, para todo  $x_1, x_2, \dots, x_d \in \mathbb{R}$*

$$F_X(x_1, x_2, \dots, x_d) = C[F_1(x_1), F_2(x_2), \dots, F_d(x_d)] \quad (3.1)$$

### 3.1. Tipos de Cópulas

Una cópula puede asumir distintas formas, dado que nuestro objetivo es cuantificar los eventos extremos, nos interesan aquellas que midan la dependencia en la cola derecha. Por un lado elegiremos la cópula Gaussiana por ser el modelo más simple, aunque no suponga dependencia en las colas. Por otro lado, elegiremos las cópulas t-Student y Gúmbel, que poseen dependencia en la cola derecha de la distribución de probabilidad, puesto que recogerán mejor las pérdidas extremas.

A continuación, introducimos el caso genérico de  $d$  factores para cada cópula, pero por simplicidad, únimamente desarrollaremos el caso bivalente.

Dentro del conjunto de funciones cópula, se pueden agrupar por familias que cumplan una serie de particularidades:

#### 3.1.1. Cópulas elípticas

Reciben este nombre porque estas cópulas provienen de distribuciones elípticas. La dependencia en este tipo de distribuciones está representada por relaciones de dependencia simétricas, siendo una herramienta poco útil para describir la dependencia en los casos que la distribución es no-Normal. Algunos ejemplos son:

---

<sup>5</sup>En el caso continuo, esta cópula será única.

### Cópula Gaussiana:

La cópula Gaussiana es una función de dependencia que deriva de una distribución Normal estándar multivariante  $\Phi_d$ , con media cero y matriz de correlación  $\Sigma$ . Lo que implica que sus distribuciones marginales normales con la cópula Gaussiana formarán distribuciones normales multivariadas:

$$C_{Gauss}(u_1, u_2, \dots, u_d; \Sigma) = \Phi_d [\Phi^{-1}(u_1), \Phi^{-1}(u_2), \dots, \Phi^{-1}(u_d)], \quad (3.2)$$

donde  $\Sigma$  denota una matriz de correlaciones cuya diagonal es unitaria, y  $\Phi^{-1}$  representa la inversa de la distribución normal univariante estándar.

Para el caso bivalente,  $d=2$ :

- La función de distribución será:

$$\begin{aligned} C_\rho(u_1, u_2; \rho) &= \Phi_2 [\Phi^{-1}(u_1), \Phi^{-1}(u_2)] \\ &= \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{1}{2\pi} \frac{1}{\sqrt{1-\rho^2}} \exp\left(\frac{2t_1 t_2 - t_1^2 - t_2^2}{2(1-\rho^2)}\right) dx_1 dx_2, \end{aligned} \quad (3.3)$$

donde  $\Phi^{-1}(u_1)$  y  $\Phi^{-1}(u_2)$  son las funciones de distribución marginal que coinciden con la función de distribución Normal estándar, por definición de la cópula Gaussiana.

- Por lo que respecta a la función de densidad:

$$c_\rho(u_1, u_2; \rho) = \frac{1}{\sqrt{1-\rho^2}} \exp\left(\frac{2t_1 t_2 - t_1^2 - t_2^2}{2(1-\rho^2)}\right). \quad (3.4)$$

### Cópula t-Student:

La cópula t-Student, como en el caso de la cópula Gaussiana, coincide con su distribución estándar. Es decir, contiene la estructura de dependencia asociada a la distribución t-Student multivariante:

$$C_{Stud}(u_1, u_2, \dots, u_d; \Sigma, \nu) = T_{d,\nu} [t_\nu^{-1}(u_1), t_\nu^{-1}(u_2), \dots, t_\nu^{-1}(u_d)], \quad (3.5)$$

siendo  $T_{d,\nu}$  la función de distribución t-Student multivariante, con matriz de correlación  $\Sigma$  y con  $\nu$  grados de libertad. Por otro lado,  $t_\nu^{-1}$  es la inversa de la función t-Student univariante con  $\nu$  grados de libertad.

Para el caso bivalente,  $d=2$ :

- La función de distribución será:

$$\begin{aligned} C_\rho(u_1, u_2; \rho, \nu) &= T_2 [t_\nu^{-1}(u_1), t_\nu^{-1}(u_2)] \\ &= \int_{-\infty}^{t_\nu^{-1}(u_1)} \int_{-\infty}^{t_\nu^{-1}(u_2)} \frac{\Gamma\left(\frac{\nu+2}{\nu}\right)}{\Gamma\left(\frac{\nu+2}{\nu}\right) 2\pi\sqrt{1-\rho^2}} \exp\left(1 + \frac{2t_1t_2 - t_1^2 - t_2^2}{\nu(1-\rho^2)}\right) dt_1 dt_2, \end{aligned} \quad (3.6)$$

donde  $t_\nu^{-1}(u_1)$  y  $t_\nu^{-1}(u_2)$  son las funciones de distribución marginal que coinciden con la función de distribución t-Student estándar, por definición de la cópula t-Student.

- La función de densidad será:

$$c_{\nu,\rho}(u_1, u_2; \rho) = \frac{\Gamma\left(\frac{\nu+2}{\nu}\right)}{\Gamma\left(\frac{\nu+2}{\nu}\right) 2\pi\sqrt{1-\rho^2}} \exp\left(1 + \frac{2t_1t_2 - t_1^2 - t_2^2}{\nu(1-\rho^2)}\right). \quad (3.7)$$

Cabe destacar que tanto la cópula Gaussiana como la cópula t-Student son *cóputas implícitas*, es decir, se obtienen a partir de la distribución conocida, puesto que su forma funcional coincide con dicha función de distribución.

### 3.1.2. Cópula arquimediana

En este tipo de cópulas existe una función que se denomina *generador de la cópula*,  $\varphi$ , que las permite construir fácilmente. Por lo tanto, van a permitir recoger muchos tipos de estructuras de asociación. Teniendo en cuenta que en la mayoría de casos los datos no están normalmente distribuidos, esta familia puede ser utilizada para los casos asimétricos.

#### Cópula Gúmbel:

La cópula Gúmbel es adecuada para datos que presenten asimetría a la derecha, es decir que haya una dependencia positiva. Si utilizamos la función generadora  $\varphi(u) = (-\ln u)^\theta$ ,

podemos representar la cópula Gumbel de la siguiente manera:

$$C_{Gumb}(u_1, u_2, \dots, u_d; \theta) = \exp \left\{ - \left[ \sum_{i=1}^n (-\ln u_i)^\theta \right]^{1/\theta} \right\}, \quad \theta \in [1, \infty). \quad (3.8)$$

Para el caso bivalente, d=2:

- La función de distribución será:

$$C_\theta(u_1, u_2) = \exp \left\{ - [(-\ln u_1)^\theta + (-\ln u_2)^\theta]^{1/\theta} \right\}. \quad (3.9)$$

En la cópula Gumbel, la dependencia es perfecta cuando  $\theta \rightarrow +\infty$  y coincide con la cópula de independencia cuando  $\theta = 1$ .

- Por lo que la función de densidad de la cópula Gumbel:

$$c_\theta(u_1, u_2) = C_\theta(u_1, u_2) u_1^{-1} u_2^{-1} \left\{ [(-\ln(u_1))^\theta + (-\ln(u_2))^\theta]^{-2+\frac{2}{\theta}} \right\} [\ln(u_1)\ln(u_2)]^{\theta-1} \left[ 1 + (\theta + 1) [(-\ln(u_1))^\theta + (-\ln(u_2))^\theta]^{-\frac{1}{\theta}} \right], \quad (3.10)$$

A diferencia de las cópulas anteriormente comentadas, la cópula Gumbel es una *cópula explícita*, es decir, posee una forma funcional propia, además de sencilla.

A modo resumen:

Tabla 2: Clasificación cópulas

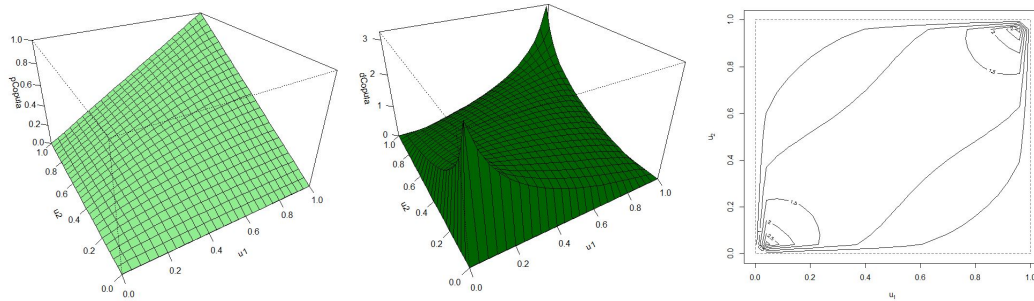
Nombre cópula	Clasificación
Gaussiana	Implícita y Elíptica
t-Student	Implícita y Elíptica
Gumbel	Explícita y Arquemediana

*Fuente: Elaboración propia*



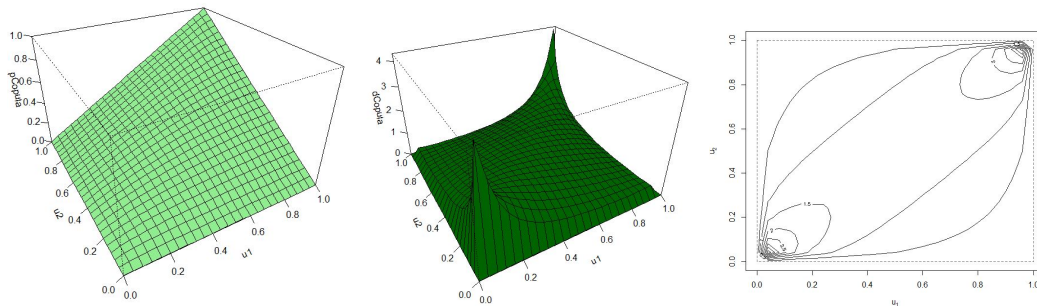
### Gráficos de las cópulas Gaussiana, t-Student y Gúmbel:

Figura 3: Función de distribución, densidad y líneas de nivel de la cópula Gaussiana,  $\rho=0.5$



Fuente: Elaboración propia

Figura 4: Función de distribución, densidad y líneas de nivel de la cópula t-Student,  $\rho=0.5$ ,  $\nu = 4$



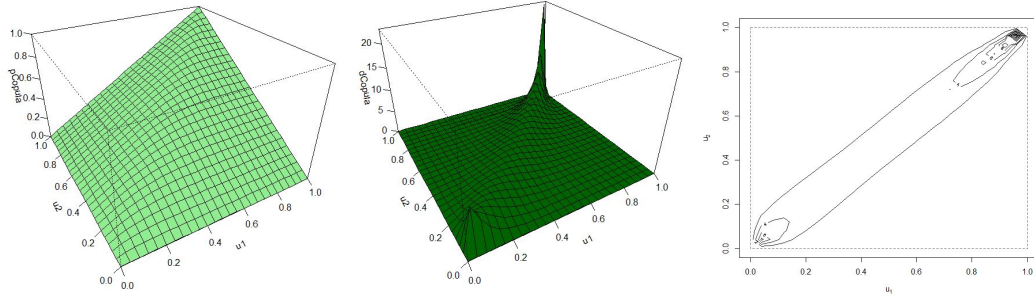
Fuente: Elaboración propia

En las Figuras 3 y 4, observamos que la cópula Gaussiana y la t-Student son simétricas,  $C(u_1, u_2) = C(u_2, u_1)$ . También comprobamos como la cópula t-Student tiene densidad superior en los valores extremos en comparación con la cópula Gaussiana, en otras palabras, la función de densidad de la cópula t-Student presenta más masa de probabilidad en las partes superior e inferior de la distribución.

Además, la cópula t-Student presenta dependencia en los cuantiles cercanos a 1, de modo que  $\lambda_u = \lambda_l \geq 0$ , a diferencia de la Gaussiana cuya dependencia en las colas es nula,

$$\lambda_u = \lambda_l = 0.$$

Figura 5: Función de distribución, densidad y líneas de nivel de la cópula Gúmbel,  $\theta=4$



Fuente: Elaboración propia

En Figura 5, podemos observar la asimetría de la cópula Gúmbel, en este caso alcanza la densidad máxima cuando ambos argumentos tienden a 1, por ello hay dependencia en la cola de la derecha. Por tanto, sí existe dependencia en las colas, pero obviamente sólo en la de la derecha:  $\lambda_u = 2 - 2^{\frac{1}{\theta}}$  y  $\lambda_l = 0$ .

A la cópula Gúmbel y a la cópula t-Student también se las conoce como cópulas de valor extremo.

## 3.2. Medidas de dependencia

La palabra “correlación” se utiliza para describir la dependencia que hay entre dos variables aleatorias. Las cópulas han proporcionado formas alternativas, al coeficiente lineal de Pearson, de definir medidas de dependencia.

### 3.2.1. Coeficiente lineal de Pearson:

El coeficiente lineal de Pearson mide la dependencia lineal entre dos variables. Formalmente sería:

$$\rho_{xy} = \frac{\sum_{i=1}^n (L_{xi} - \hat{\mu}_x)(L_{yi} - \hat{\mu}_y)}{\sqrt{\sum_{i=1}^n (L_{xi} - \hat{\mu}_x)^2 \sum_{i=1}^n (L_{yi} - \hat{\mu}_y)^2}} \quad (3.11)$$

donde  $\hat{\mu}_x = \frac{\sum_{i=1}^n L_{xi}}{n}$  y  $\hat{\mu}_y = \frac{\sum_{i=1}^n L_{yi}}{n}$  son las medias muestrales de cada pérdida. El coeficiente de correlación  $\rho_{xy}$  nos proporciona la dependencia lineal que existe entre las pérdidas. Esto implica que si existe relación no lineal entre las pérdidas no lo va a tener en cuenta y por lo tanto no devolverá un resultado correcto. Cuando se obtienen los valores extremos  $\pm 1$  indica la máxima dependencia. Si el signo es positivo, significa que si una variable aumenta la otra variable también, mientras que si el signo es negativo, se moverán en sentido opuesto. Cuando  $\rho_{xy}$  tiene valor nulo, equivale a que hay independencia entre las variables.

### 3.2.2. Medidas de dependencia utilizando cópulas

Así pues, teniendo en cuenta las limitaciones de las medidas de relación lineal, definiremos otras medidas de asociación que valoran comportamientos que no son necesariamente normales ni lineales. Como la Tau de Kendall, en la que se basan la mayoría de estimadores, o los coeficientes de dependencia de las colas, siendo muy importantes en la cuantificación del riesgo, sobre todo para la cola derecha.

#### **Tau de Kendall - Concordancia y Discordancia:**

Se define como la probabilidad de concordancia menos la probabilidad de discordancia. Dada una muestra de  $n$  pérdidas observadas  $(L_{xi}, L_{yi})$ , se estimará la  $\tau$  de Kendall como se muestra:

$$\hat{\tau} = \frac{c - d}{c + d} = \frac{c - d}{\frac{1}{2}n(n - 1)}, \quad (3.12)$$

donde  $c$  será el número de pares concordantes y  $d$  será el número de pares discordantes.

Es decir, para cada par de observaciones consecutivas  $(L_{xi}, L_{yi})$  y  $(L_{xj}, L_{yj})$  se consideran que son concordantes en el caso que  $(L_{xi} - L_{xj})(L_{yi} - L_{yj}) > 0$  y se consideran discordantes en caso que se cumpla lo contrario,  $(L_{xi} - L_{xj})(L_{yi} - L_{yj}) < 0$ . Cuando  $L_{xi} = L_{xj}$  y  $L_{yi} = L_{yj}$  las observaciones no serán ni concordantes ni discordantes.

Esta medida de asociación puede deducirse a partir de la cópula:

$$\tau = 4 \int_0^1 \int_0^1 C_\theta(u_x, u_y) dC_\theta(u_x, u_y) - 1. \quad (3.13)$$

Cabe destacar, que ni la cópula Gaussiana ni la t-Student tienen expresiones cerradas para la  $\tau$  de Kendall. Consiguientemente, se deberán aproximar a partir de sus expresiones en función de la cópula. Por otro lado, es interesante mencionar que la cópula de Gumbel, permite deducir el coeficiente de correlación de Kendall en función de su parámetro  $\theta$ :

$$\tau(\theta) = 1 - \frac{1}{\theta}. \quad (3.14)$$

### Coefficiente de dependencia de la cola:

Este concepto hace referencia al nivel de relación en la parte superior e inferior de una distribución multivariante. Será más relevante cuando trabajemos con estructura de dependencia entre valores extremos. A partir del coeficiente de dependencia de la cola, se pueden deducir dos coeficientes, el de la cola derecha de la distribución (upper tail dependence) y el de la cola izquierda (lower tail dependence), a los que denominamos  $\lambda_u$  y  $\lambda_l$ , respectivamente.

Estos coeficientes se pueden deducir a partir de la cópula. En primer lugar, el coeficiente  $\lambda_u$ , será:

$$\lambda_u = \lim_{u \rightarrow 1^-} \frac{\bar{C}_\theta(u, u)}{1 - u}, \quad (3.15)$$

donde  $\bar{C}_\theta(u, u) = 1 - C_\theta(u, u)$  es la cópula de supervivencia e indicamos que el límite tiende a 1 por la derecha, es decir, que las pérdidas tienden a  $+\infty$  por la izquierda. Por lo tanto, el coeficiente de dependencia de la cola izquierda,  $\lambda_l$ , será:

$$\lambda_l = \lim_{u \rightarrow 0^+} \frac{C_\theta(u, u)}{u}, \quad (3.16)$$

en este caso, los argumentos del límite tienden a  $-\infty$  por la derecha.

## 3.3. Estimación de los Parámetros y Criterios de Selección de la Cópula

### 3.3.1. Estimación de los parámetros

Existen varios métodos estadísticos para estimar parámetros, aunque bajo este contexto se considerará el método de pseudo-máximo verosimilitud por ser más fácil de optimizar.

## Método Máxima Verosimilitud

En este método, se maximiza de manera conjunta el logaritmo de la función de verosimilitud asociada a la cópula y a las marginales.

Por lo tanto, si  $\hat{\theta}$  es el parámetro estimado de la cópula y  $\hat{\beta}_1$  y  $\hat{\beta}_2$  los parámetros estimados de las distribuciones marginales, el logaritmo de su función de verosimilitud será:

$$\mathcal{L}(\hat{\theta}, \hat{\beta}_1, \hat{\beta}_2) = \sum_{i=1}^n \ln \{c_{\hat{\theta}}(U_{i1}, U_{i2})\} + \ln \{f_1(L_{i1}; \hat{\beta}_1)\} + \ln \{f_2(L_{i2}; \hat{\beta}_2)\}, \quad (3.17)$$

donde  $c_{\hat{\theta}}$  es la densidad de la cópula,  $U_{i1} = F_1(L_{i1}; \hat{\beta}_1)$  y  $U_{i2} = F_2(L_{i2}; \hat{\beta}_1)$  (funciones de distribución de las marginales) y  $f_1, f_2$ , las funciones de densidad de las marginales. Siendo la función de densidad la derivada de la función de distribución. No suele aplicarse este método debido a que es difícil encontrar unos parámetros de cópula y marginales que optimizen  $\mathcal{L}(\hat{\theta}, \hat{\beta}_1, \hat{\beta}_2)$ .

## Método Pseudo-máxima Verosimilitud

Se le denomina Pseudo-máxima Verosimilitud, puesto que el resultado obtenido no maximizará la verosimilitud total. A diferencia del método anterior, se maximizará la verosimilitud de manera parcial, es decir, por un lado la verosimilitud de la cópula y por otro lado la verosimilitud de las marginales.

En primer lugar se generarán pseudo-observaciones (distribuciones univariantes)  $(\tilde{U}_{i1}, \tilde{U}_{i2})$ ,  $i = 1, \dots, n$ , que se determinarán mediante las funciones de distribución empíricas corregidas:

$$\tilde{U}_{i1} = \frac{n}{n+1} \sum_{j=1}^n I(L_{j1} \leq L_{i1}) \quad y \quad \tilde{U}_{i2} = \frac{n}{n+1} \sum_{j=1}^n I(L_{j2} \leq L_{i2}), \quad (3.18)$$

Entonces, con las pseudo-observaciones obtenemos los argumentos de la función pseudo-verosimilitud a maximizar respecto al parámetro de la cópula, es decir, se debe maximizar la siguiente expresión:

$$\mathcal{L}(\hat{\theta}) = \sum_{i=1}^n \ln \{c_{\hat{\theta}}(\tilde{U}_{i1}, \tilde{U}_{i2})\}, \quad (3.19)$$

La metodología referente a las cópulas se ha obtenido de los artículos de Bolancé *et al.* (2015a), Torres y Olarte (2009a) y Frees y Valdez (1998a).

### 3.3.2. Criterios de Selección

Es fundamental determinar la cópula que mejor ajuste a la estructura de dependencia para el conjunto de distribuciones. Una manera es mediante los diferentes criterios de información y seleccionar la cópula que mejor corresponde a la distribución de las observaciones. Para ello, utilizaremos dos tipos de criterios de selección:

#### AIC- Akaike Information Criterion

Se trata de una medida de calidad de ajuste de un modelo con respecto a los datos utilizados. Esta medida tiene en cuenta el número de parámetros utilizados en el modelo. Se define como:

$$AIC = -2\log(\hat{\theta}) + 2k, \quad (3.20)$$

siendo  $\theta$  el vector de parámetros,  $\log(\hat{\theta})$  el logaritmo de la función verosimilitud que maximiza  $\theta$  y  $k$  es el número de parámetros estimados.

#### BIC- Bayesian Information Criterion

También es una medida de calidad de ajuste de un modelo para los datos utilizados. Esta medida, a diferencia del AIC, tiene en cuenta el tamaño muestral. Se define como:

$$BIC = -2\log(\hat{\theta}) + k\log(n), \quad (3.21)$$

siendo  $\theta$  el vector de parámetros,  $\log(\hat{\theta})$  el logaritmo de la función verosimilitud que maximiza  $\theta$ ,  $k$  es el número de parámetros estimados y  $n$  es el tamaño muestral.

En ambas medidas, cuanto menor es el valor, mejor será el ajuste.

### 3.4. Medidas de riesgo

Existen diferentes formas para cuantificar el riesgo de las pérdidas obtenidas. Las medidas de riesgo, en términos generales, nos asocian el valor de la función de la variable aleatoria pérdida,  $L$ , con un número real que nos permite comparar varios resultados fácilmente. La más conocida probablemente sea el *Value at Risk* (VaR) (Veáse Artzner *et al.*, 1999a).

**Definición 3.2 (value at risk)** Dado un nivel de confianza  $\alpha \in (0, 1)$ , el VaR de una cartera de pérdida  $L$  con un nivel de confianza  $\alpha$ , será el valor más pequeño  $l$ , con una probabilidad no mayor a  $1 - \alpha$  de que la pérdida  $L$  exceda  $l$ .

$$VaR_\alpha = VaR_\alpha(L) = \inf \{l \in \mathbb{R} : P(L > l) \leq 1 - \alpha\} = \inf \{l \in \mathbb{R} : F_l \geq \alpha\} \quad (3.22)$$

Siendo  $L$  una variable aleatoria que representa la pérdida total en una situación de riesgo y  $F_l$  su función de distribución. Por lo tanto, el VaR es el cuantil  $(1 - \alpha)$  de la función de distribución de la variable pérdida o lo que es lo mismo la probabilidad de tener una pérdida superior a la determinada por el VaR será  $(1 - \alpha)$ .

Otra medida muy conocida es *Tail Value at Risk* (TVaR):

**Definición 3.3 (tail value at risk)** pérdida esperada dado que ésta ha excedido el percentil  $\alpha$  de la distribución de la pérdida  $L$ .

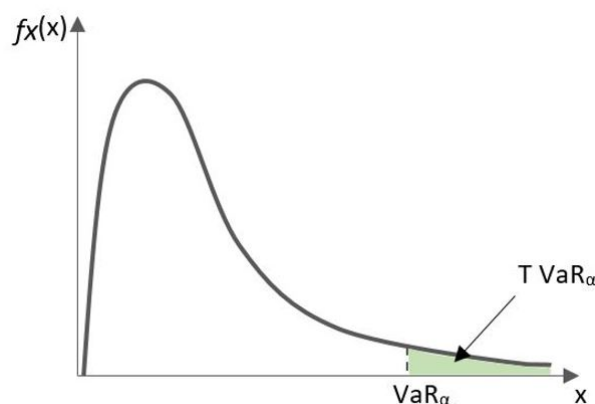
$$TVaR_\alpha(L) = \frac{1}{1 - \alpha} \int_\alpha^1 VaR_u(L) du \quad (3.23)$$

Para variables aleatorias continuas TVaR estará dado por:

$$TVaR_\alpha(L) = E[L|L > VaR_\alpha(L)] \quad (3.24)$$

Esta medida es muy interesante puesto que nos da información de colas pesadas. En la Figura 6 representamos el concepto de VaR y TVaR utilizando la función de densidad de un riesgo.

Figura 6: Gráfico VaR-TVaR



Fuente: Elaboración propia

### 3.5. Estimación del riesgo integral del asegurado

En primer lugar modelizamos la probabilidad de que exista siniestro, en base a distintas variables explicativas. Tenemos  $d$  variables dependientes dicotómicas, que tomarán valor 1 cuando se declare el siniestro y valor 0 en caso contrario. Nuestro objetivo será modelizar las tres líneas de negocio para comprobar si están correlacionadas, de esta forma, tendrá sentido aplicar un Probit Multivariante.

En segundo lugar tendremos que modelizar los costes que se han pagado al asegurado en compensación al accidente (severidad) para así cuantificar el riesgo, mediante las cópulas. Para ello, utilizaremos el método de simulación de Monte Carlo. A partir la forma que nos muestra el histograma de los costes ajustaremos diferentes distribuciones que consideramos se adecuan mejor a la muestra. Seguidamente generaremos pérdidas asociadas a estos modelos ajustados, para poder estimar el riesgo de forma empírica.

El procedimiento que aplicaremos para generar valores a partir de una cópula será el que sigue:

1. Transformación de las variables originales a variables uniformes, a las que llamamos pseudo-observaciones  $(\tilde{U}_{i1}, \tilde{U}_{i2}, \dots, \tilde{U}_{id})$ ,  $i = 1, \dots, n$ ; de la  $d$ -variante cópula, usando la función de distribución empírica:

$$\tilde{U}_{id} = \frac{n}{n+1} \sum_{j=1}^n I(L_{jd} < L_{id}). \quad (3.25)$$

2. Estimación de los parámetros de las diferentes cópulas mediante el método de pseudo-máxima verosimilitud. Deberemos maximizar la siguiente expresión:

$$\mathcal{L}(\tilde{\theta}) = \sum_{i=1}^n \ln \left\{ c_{\tilde{\theta}}(\tilde{U}_{i1}, \tilde{U}_{i2}, \dots, \tilde{U}_{id}) \right\}. \quad (3.26)$$

3. Elegiremos la cópula bajo los criterios de Akaike Information Criterion,  $AIC$  y Bayesian Information Criterion  $BIC$ . Se seleccionará aquella que muestre menor  $AIC$  y  $BIC$ .
4. A partir de la cópula elegida, simularemos  $(\tilde{U}_{r1}, \tilde{U}_{r2}, \dots, \tilde{U}_{rd})$ ,  $r = 1, \dots, R$ . Siendo  $R$  el número de réplicas que aplicamos, valor que deberá ser suficientemente grande puesto que son números aleatorios y se asume que son todos los casos posibles que pueden ocurrir. Por ello, cuanto mayor sea, más probabilidad de representar todos los escenarios posibles. Tomaremos  $R=1,000,000$ .



5. Utilizando la inversa de la función de distribución, calcularemos las pérdidas simuladas:  $\tilde{L}_{r1} = F_1^{-1}(\tilde{U}_{r1})$ ,  $\tilde{L}_{r2} = F_1^{-1}(\tilde{U}_{r2})$ , ...,  $\tilde{L}_{rd} = F_1^{-1}(\tilde{U}_{rd})$ .
6. Obtendremos el cuantil  $q_\alpha$  para  $\alpha = 0.99, 0.995, 0.999$ , de la suma de los costes simulados. Para facilitar la lectura de los datos, se mostrarán en miles de euros.

Por último, una vez hemos obtenido la probabilidad y hemos cuantificado el riesgo obtendremos la pérdida integrada. Para el caso práctico planteado en este trabajo será tal que:

$$\begin{aligned}
L = & VaR_L(C_{md}, C_{bi}, C_{home}) \times Pr(y_{md} = 1, y_{bi} = 1, y_{home} = 1) + \\
& VaR_L(C_{md}, C_{bi}) \times Pr(y_{md} = 1, y_{bi} = 1, y_{home} = 0) + \\
& VaR_L(C_{md}, C_{home}) \times Pr(y_{md} = 1, y_{bi} = 0, y_{home} = 1) + \\
& VaR_L(C_{bi}, C_{home}) \times Pr(y_{md} = 0, y_{bi} = 1, y_{home} = 1) + \\
& VaR_L(C_{md}) \times Pr(y_{md} = 1, y_{bi} = 0, y_{home} = 0) + \\
& VaR_L(C_{bi}) \times Pr(y_{md} = 0, y_{bi} = 1, y_{home} = 0) + \\
& VaR_L(C_{home}) \times Pr(y_{md} = 0, y_{bi} = 0, y_{home} = 1), \\
L = & TVaR_L(C_{md}, C_{bi}, C_{home}) \times Pr(y_{md} = 1, y_{bi} = 1, y_{home} = 1) + \\
& TVaR_L(C_{md}, C_{bi}) \times Pr(y_{md} = 1, y_{bi} = 1, y_{home} = 0) + \\
& TVaR_L(C_{md}, C_{home}) \times Pr(y_{md} = 1, y_{bi} = 0, y_{home} = 1) + \\
& TVaR_L(C_{bi}, C_{home}) \times Pr(y_{md} = 0, y_{bi} = 1, y_{home} = 1) + \\
& TVaR_L(C_{md}) \times Pr(y_{md} = 1, y_{bi} = 0, y_{home} = 0) + \\
& TVaR_L(C_{bi}) \times Pr(y_{md} = 0, y_{bi} = 1, y_{home} = 0) + \\
& TVaR_L(C_{home}) \times Pr(y_{md} = 0, y_{bi} = 0, y_{home} = 1),
\end{aligned}$$

donde la medida de riesgo será el VaR o el TVaR.

## 4. Contexto práctico

En esta sección se procederá a la aplicación de la metodología y posterior análisis de un portfolio de seguros no vida. Nuestro objetivo será calcular la pérdida total a partir de la agregación de los distintos posibles escenarios en los que se declara siniestro.

### 4.1. Descripción de los Datos

Los datos utilizados son de una compañía de seguros española e incluye información de una cartera de seguros no vida conformada por tres líneas de negocio: daños materiales, daños corporales y hogar. Recoge 161,285 observaciones entre los años 2006-2015.

En la Tabla 3 se muestran las variables de dicha base de datos:

Tabla 3: Explicación de las variables

<b>client_id</b>	Identificación cliente.
<b>client_sex</b>	Toma valor 1 si es hombre y 0 si es mujer.
<b>client_age</b>	Edad del asegurado (18 a 85 años).
<b>lic_age</b>	Antigüedad del permiso de conducir (0 a 68 años).
<b>nother</b>	Número de pólizas contratadas en la compañía además de daños materiales, daños corporales y hogar.
<b>cities2</b>	Toma valor 1 si la zona de conducción es Madrid-Barcelona y 0 en caso contrario.
<b>north</b>	Toma valor 1 si la zona de conducción es Norte de España y 0 en caso contrario.
<b>rest</b>	Toma valor 1 si la zona de conducción es el resto de España y 0 en caso contrario.
<b>cost_md</b>	Coste ramo automóvil - Daños materiales.
<b>nclaims_md</b>	Número de siniestros ramo automóvil - Daños materiales.
<b>y_md</b>	Toma valor 1 si el asegurado declara siniestros de daños materiales durante el periodo de exposición y 0 en caso contrario.
<b>cost_bi</b>	Coste ramo auto - Daños corporales.
<b>nclaims_bi</b>	Número de siniestros ramo auto - Daños corporales.
<b>y_bi</b>	Toma valor 1 si el asegurado declara siniestros de daños corporales durante el periodo de exposición y 0 en caso contrario.

<b>exposi_auto</b>	Exposición al riesgo ramo automóvil.
<b>cost_home</b>	Coste ramo hogar.
<b>nclaims_home</b>	Número de siniestros ramo hogar
<b>y_home</b>	Toma valor 1 si el asegurado declara siniestros de hogar. durante el periodo de exposición y 0 en caso contrario.
<b>exposi_home</b>	Exposición al riesgo ramo hogar.

*Fuente: Elaboración propia*

A continuación se muestran los estadísticos descriptivos de cada variable separadas en tres tablas, dos para las variables dependientes y otra para las variables explicativas:

Tabla 4: Estadísticos descriptivos para las variables dependientes - Número de siniestros

	número de siniestros (n_claims)				existencia de siniestro (y)			
	Media	Varianza	Mín.	Máx.	Media	Varianza	Mín.	Máx.
<b>Auto_md</b>	0.2574821	0.679877	0	40	0.1521964	0.1290335	0	1
<b>Auto_bi</b>	0.0352110	0.042837	0	6	0.0315342	0.0305400	0	1
<b>Hogar</b>	0.2020709	0.355604	0	23	0.1438881	0.1231851	0	1

*Fuente: Elaboración propia*

Según la Tabla 4, la categoría que tiene mayor siniestralidad es daños materiales, seguida de hogar y por último la de daños corporales que es la que menos casos se producen. Si comparamos con la variable binaria  $y$  observamos como las variables daños materiales y hogar disminuyen respecto el número de siniestros, mientras que en daños corporales no se produce una disminución tan acusada. Esto nos indica que tanto en daños materiales como en hogar, los siniestros están menos repartidos que en daños corporales y que unos pocos asegurados copan una cantidad de siniestros excesiva.

Tabla 5: Estadísticos descriptivos para las variables dependientes - Costes

	coste global				coste y=1			
	Media	Varianza	Mín.	Máx.	Media	Varianza	Mín.	Máx.
<b>Auto_md</b>	238.6432430	1097030.60	0	153177.94	882	5123739	0	153177.94
<b>Auto_bi</b>	219.6774997	18510259.02	0	1023100	6966	540092541	0	1023000
<b>Hogar</b>	63.6337449	883741.74	0	161666.26	442.2	59746470	0	161666.26

*Fuente: Elaboración propia*

En la Tabla 5, se muestra la diferencia de estadísticos descriptivos. En la parte izquierda se presenta la información de costes con todos los datos, en cambio en la parte derecha se han condicionado los costes a la existencia de siniestro. Por tanto, como se han obviado los sujetos con siniestralidad nula, se muestra la descripción real de los costes para cada ramo. Destacamos que el coste de los daños corporales es el más elevado.

Tabla 6: Estadísticos descriptivos para las variables explicativas

	población total				casos con siniestro			
	Media	Varianza	Mín.	Máx.	Media	Varianza	Mín.	Máx.
<b>client_sex</b>	0.7621416	0.1812829	0	1	0.7820167	0.1704704	0	1
<b>client_age</b>	53.1169669	167.5265024	18	85	53.2922125	149.4392223	18	85
<b>lic_age</b>	29.6650277	127.8406338	0	68	29.8623002	111.6412153	1	68
<b>nother</b>	0.2202127	0.4341144	0	23	0.2981319	0.6307354	0	23
<b>cities2</b>	0.1946182	0.1567429	0	1	0.1872834	0.1522117	0	1
<b>north</b>	0.2905540	0.2061336	0	1	0.2880936	0.2051003	0	1
<b>rest</b>	0.5148278	0.2497817	0	1	0.5246230	0.2493993	0	1

*Fuente: Elaboración propia*

En la Tabla 6, se ha hecho una comparativa de las variables independientes en términos de si se ha declarado siniestro o no. El bloque izquierdo nos muestra los datos estadísticos del global de la cartera, mientras que el bloque derecho describe el grupo de asegurados que ha sufrido como mínimo un siniestro. Cabe puntualizar que los resultados son parecidos para los dos análisis.

## 4.2. Aplicación del Probit Multivariante

Consideramos un portfolio de pólizas de seguros, con  $d = 3$  líneas de negocio. Por lo tanto, el total de posibles combinaciones equivale a  $2^d$ , como sólo tenemos en cuenta los casos en los que como mínimo se ha dado un tipo de siniestro, será  $2^d - 1$ . Entonces, para esta cartera tenemos un total de 7 combinaciones:

M	1	2	3	4	5	6	7
Combinaciones	(1,1,1)	(1,1,0)	(1,0,1)	(0,1,1)	(1,0,0)	(0,1,0)	(0,0,1)

M=1: se declaran los tres tipos de siniestros.

M=2, M=3, M=4 : se declaran los dos tipos de siniestros y hay un tipo de siniestro que no se declara.

M=5, M=6, M=7: se declara un tipo de siniestro y hay dos tipos de siniestro que no se declaran.

Cabe destacar que la duración típica de las coberturas de automóvil y hogar es de 1 año. Como los seguros no suelen contrarse al inicio del año, este tipo de datos tienen una variable para tratar la duración de la cobertura como una fracción del año, se conoce como exposición y es el periodo que un asegurado ha estado cubierto por el seguro. Para el cálculo lo trataremos como *offset*, es decir como una variable adicional que estará limitada a 1. Como aplicamos un probit multivariante, el *offset* debe tener la misma escala que el predictor lineal, por lo que hemos tenido que aplicar el cuantil de la normal estándar a la exposición.

En el Anexo II se encuentra el código *SAS* con el que hemos realizado los cálculos.

Tabla 7: Matriz correlación entre los diferentes riesgos

	Daños materiales	Daños corporales	Hogar
Daños materiales	1	0.796210 (***)	0.052072(***)
Daños corporales	0.796210(***)	1	0.053115(***)
Hogar	0.052072(***)	0.053115(***)	1

Nivel significación  $\alpha = 5\%$ (\*\*),  $\alpha = 1\%$ (\*\*\*)

La Tabla 7, corresponde a la matriz de correlaciones. En base a los resultados concluimos que tiene sentido aplicar un modelo multivariante, puesto que las variables son dependientes y todas las correlaciones son significativas, por consiguiente aportan información a la hora de estimar el modelo. Observamos que incluso las correlaciones de daños materiales-hogar que presentan valores muy bajos son significativas. Cabe puntualizar, que la correlación más elevada se da con los daños materiales y daños corporales, esto es lógico ya que relacionan dos riesgos asociados al mismo seguro de automóvil. Determinamos que los resultados que se exponen a continuación son correctos, porque son los que están calculados de forma más eficiente y esta correlación afecta a la significación de los parámetros. Si no se tuviera en cuenta se estarían subestimando sus varianzas y por tanto tenderíamos a concluir que todas las variables tienen efecto significativo.

Tabla 8: Estimación de los parámetros para cada línea de negocio

<b>Variable</b>	<b>Estimación Daños materiales</b>	<b>Estimación Daños corporales</b>	<b>Estimación Hogar</b>
Intercept	0.969531 <sup>(***)</sup>	0.081659 <sup>(**)</sup>	1.019694 <sup>(***)</sup>
Sexo del asegurado: Hombre	-0.140253 <sup>(***)</sup>	-0.107191 <sup>(***)</sup>	0.0137588
Edad del asegurado	0.030160	0.118761	0.145599 <sup>(***)</sup>
Antigüedad permiso conducir	-0.743127 <sup>(***)</sup>	-1.0022636 <sup>(***)</sup>	-
Otras pólizas	0.020086 <sup>(***)</sup>	-0.048717 <sup>(***)</sup>	0.026245 <sup>(***)</sup>
Zona conducción:Madrid-Barcelona	-0.031336 <sup>(***)</sup>	0.049139 <sup>(***)</sup>	0.013955
Zona conducción: Norte España	-0.000466	-0.032154 <sup>(**)</sup>	-0.040872 <sup>(***)</sup>
Zona conducción: Resto España	0	0	0

Nivel significación  $\alpha = 5\%$ (\*\*),  $\alpha = 1\%$ (\*\*\*)

*Fuente: Elaboración propia*

La Tabla 8, presenta la estimación de las covariadas y su nivel de significación para cada una de las líneas de negocio. Podemos ver que en su mayoría son significativas para un  $\alpha = 1\%$  o  $\alpha = 5\%$ , por lo tanto, estas variables aportarán información para determinar la probabilidad de ocurrencia de un siniestro.

En los ramos de daños materiales y daños corporales, son significativas la constante y los parámetros asociados a las variables explicativas sexo asegurado, antigüedad del permiso de conducir, otras pólizas y zona de conducción en las ciudades Madrid y Barcelona,

siendo para daños corporales también significativa la zona de conducción en el norte de España. Por su parte, en el ramo hogar, serán significativos la constante y los parámetros de las variables edad asegurado, otras pólizas y zona de conducción norte de España.

Como es un modelo no lineal, los parámetros no son los efectos marginales, así pues no podemos interpretar la estimación de los parámetros de forma directa sobre los valores de las variables, aunque sí podemos sacar conclusiones por sus signos. De forma que si el parámetro es positivo, provoca que la probabilidad de declarar el siniestro sea mayor y si es negativo provocará lo opuesto. Por ejemplo, la constante en los tres ramos implicará que haya una mayor probabilidad de declarar siniestro, en cambio si el asegurado conduce por el norte de España, genera una menor probabilidad. Si el asegurado es hombre, para daños corporales y materiales, disminuirá la probabilidad a diferencia del ramo hogar que ocurre lo opuesto. Por su parte, cuanto mayor sea la edad del asegurado más probabilidad de declarar siniestro obtendremos. En cambio, con la antigüedad del permiso de conducir, sucede el efecto contrario.

### 4.3. Perfiles

Para poder representar el riesgo y calcular la probabilidad de que un individuo declare siniestro se han considerado cuatro perfiles. En función de las diferentes variables explicativas se han asociado a cada perfil unos valores que se exponen en la siguiente tabla:

Tabla 9: Perfiles

<b>Variabes</b>	<b>PERFIL 1</b>	<b>PERFIL 2</b>	<b>PERFIL 3</b>	<b>PERFIL 4</b>
$x_1$ : Sexo asegurado	Mujer: 0	Mujer: 0	Hombre: 1	Hombre: 1
$x_2$ : Edad asegurado	37	19	25	60
$x_3$ : Antigüedad permiso de conducir	18	1	3	41
$x_4$ : Otras pólizas	1	0	3	1
$x_5$ : Zona conducción:Madrid-Barcelona	1	0	0	1
$x_6$ : Zona conducción: Norte España	0	0	1	0
$x_7$ : Zona conducción: Resto España	0	1	0	0

*Fuente: Elaboración propia*

Para los perfiles hemos escogido dos hombres y dos mujeres, con edades y su respectiva antigüedad de carnet distintos para cada uno de ellos. Esta cartera se centra en tres zonas

geográficas, por ello se han elegido dos perfiles de Barcelona/Madrid con diferente sexo, un individuo de la zona norte de España y por último uno que su zona es cualquier parte de España exceptuando las mencionadas anteriormente. Además, también se ha tenido en cuenta si el asegurado dispone de pólizas en otros ramos de la entidad, como la media es próxima a cero, hemos considerado dos individuos que tienen otra póliza, otro que no tiene ninguna más y el restante posee tres pólizas más.

Sea  $y_i$  una variable aleatoria binaria, que corresponde a 1 cuando el asegurador declara el siniestro y 0 en caso contrario. En la Tabla 10, observamos los resultados de calcular cada una de las combinaciones de probabilidad para cada perfil, sustituyendo en cada caso el parámetro correspondiente. Para llevar a cabo la calibración se ha utilizado el código R, que se encuentra en el Anexo IV.

Tabla 10: Resultados del cálculo de probabilidad

	<b>PERFIL 1</b>	<b>PERFIL 2</b>	<b>PERFIL 3</b>	<b>PERFIL 4</b>
$Pr(y_{md} = 1, y_{bi} = 1, y_{home} = 1)$	0.4115549	0.4252129	0.3633445	0.3092828
$Pr(y_{md} = 1, y_{bi} = 1, y_{home} = 0)$	0.05771433	0.06779071	0.05106907	0.03877884
$Pr(y_{md} = 1, y_{bi} = 0, y_{home} = 1)$	0.2839457	0.2737775	0.3402907	0.3104827
$Pr(y_{md} = 0, y_{bi} = 1, y_{home} = 1)$	0.00782554	0.007295146	0.004047292	0.008417038
$Pr(y_{md} = 1, y_{bi} = 0, y_{home} = 0)$	0.04510453	0.04942243	0.05431758	0.04390415
$Pr(y_{md} = 0, y_{bi} = 1, y_{home} = 0)$	0.001220068	0.001293901	0.0006334073	0.001162995
$Pr(y_{md} = 0, y_{bi} = 0, y_{home} = 1)$	0.1639832	0.1462587	0.1583129	0.2490094

*Fuente: Elaboración propia*

#### 4.4. Aplicación de las cópulas

Partiendo de nuestra base de datos, para la modelización de la cuantía de la cartera deberemos considerar los asegurados que hayan declarado siniestro(s). Es decir, que tengan un coste mayor que cero. Nos interesan sólo estas observaciones puesto que, cuantificaremos los casos en que la variable dependiente, coste de siniestro, esté condicionada a que el asegurado declare siniestros:  $cost_i = cost_i | n\_claim > 0$ .

Nuestro objetivo consiste en calcular el coste de todas las posibles combinaciones cuando se declare cada uno de los tres tipos de siniestros obteniendo los siguientes valores del riesgo asociado a vectores multivariantes o variables univariantes:



$$VaR_L(C_{md}, C_{bi}, C_{home}), VaR_L(C_{md}, C_{bi}), VaR_L(C_{md}, C_{bi}), VaR_L(C_{bi}, C_{home}), VaR_L(C_{md}), VaR_L(C_{bi}), VaR_L(C_{home})$$

Para modelizar los costes, hemos creado un código R del Anexo III.

#### 4.4.1. Análisis del caso trivariante: Daños materiales - Daños corporales - Hogar

Hemos considerado los asegurados que han declarado siniestros en las tres coberturas. Trabajaremos con 735 observaciones que cumplen esta condición.

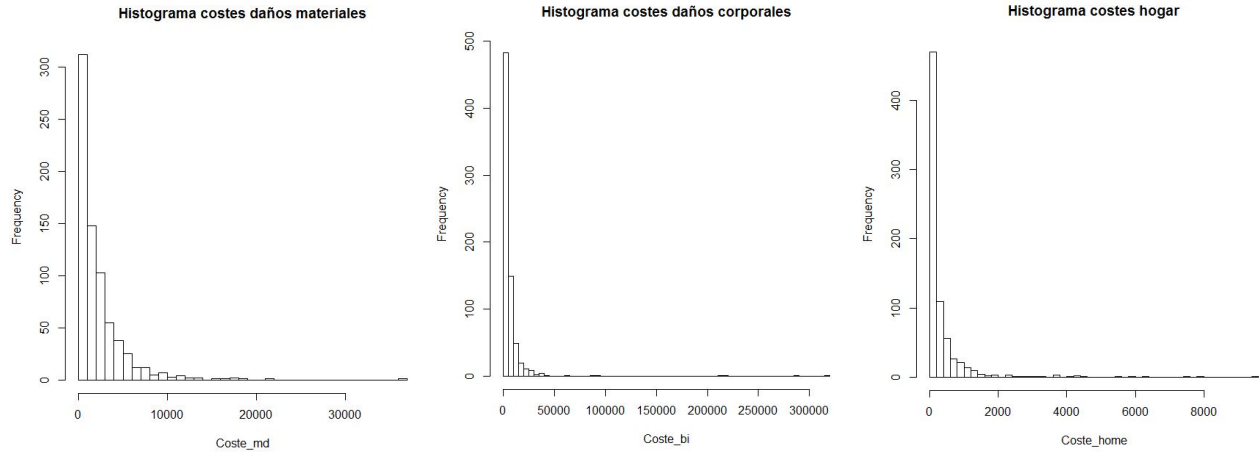
Tabla 11: Estadísticos descriptivos caso trivariante: Daños materiales - Daños corporales - Hogar

VARIABLES	Media	Desv.Estándard	Asimetría	Mínimo	Máximo	Mediana
cost_md	2426	2873.957	4.42886	0	36511	1676
cost_bi	6809	20442.82	11.36685	0	3163180	3340
cost_home	347.5	831.9946	6.075778	0	9578.9	130.0

*Fuente: Elaboración propia*

En la Tabla 11, resumimos los estadísticos descriptivos de la muestra para los costes de daños materiales (*cost\_md*), costes de daños corporales (*cost\_bi*) y costes de hogar (*cost\_home*). Las variables mencionadas siempre son positivas o cero. Vemos que hay mucha diferencia entre el valor máximo y mínimo, destacando muy por encima el valor de los costes de daños corporales en comparación a los otros dos. La asimetría en los tres casos es positiva.

Figura 7: Histogramas: Daños materiales, Daños corporales, Hogar



*Fuente: Elaboración propia*

Por lo que respecta a la dispersión, si nos fijamos en la escala del coste de los histogramas, vemos como en los `cost.bi` los valores presentan una mayor dispersión, seguidos de los `cost.md` y por último los `cost.home`, esto también se refleja en la Tabla 11.

Por su parte, en la Figura 7, se representan las formas de las distribuciones asociadas con las variables. Nos resulta muy útil, ya que nos permite analizar visualmente qué distribuciones pueden ajustar mejor. Se vé claramente que necesitaremos distribuciones para datos positivos. Por otro lado, como nuestro interés es analizar la parte de la cola derecha, ya que calcularemos el VaR con diferentes niveles de significación para ambas medidas de riesgo, se han elegido las siguientes cópulas: Gaussiana, t-Student y Gúmbel.

Tabla 12: Estimación parámetros caso trivariante: Daños materiales - Daños corporales - Hogar

<b>Cóputas</b>	<b>Parámetros iniciales</b>	<b>Parámetros finales</b>	<b>Loglikelihood</b>
Gaussian	$\rho_1 = 0,13074732$	$\rho_1 = 0,153601$	16.32
	$\rho_2 = 0,09795428$	$\rho_2 = 0,202823$	
	$\rho_3 = -0,02257349$	$\rho_3 = -0,006089$	
t-Student	$\rho_1 = 0,13074732$	$\rho_1 = 0,12881$	20.3
	$\rho_2 = 0,09795428$	$\rho_2 = 0,18843$	
	$\rho_3 = -0,02257349$	$\rho_3 = -0,01045$	
	df=4	df= 10.74397	
Gúmbel	1.040593	1.047	5.613

*Fuente: Elaboración propia*

En la Tabla 12 se muestra la estimación de los parámetros de las cóputas estudiadas. La cóputa t-Student posee dos tipos de parámetros, dependencia y grados de libertad, a diferencia del resto de las cóputas que presentan un único parámetro.

Por otro lado, la cóputa que tiene una mayor verosimilitud es la t-Student. Aunque para la elección de la cóputa se han usado los criterios de información AIC y BIC.

Los correspondientes resultados se muestran en la Tabla 13 concluyendo que la cóputa t-Student es la que mejor refleja la estructura de dependencia de nuestros datos.

Tabla 13: Elección cóputa caso trivariante: Daños materiales - Daños corporales - Hogar

<b>Cóputa</b>	<b>AIC</b>	<b>BIC</b>
Gaussiana	-26.63899	-12.83937
t-Student	-32.59429	-14.1948
Gúmbel	-9.225177	-4.625306

*Fuente: Elaboración propia*

En la Tabla 14 observamos la dependencia que hay en las colas de cada cóputa. La cóputa Gaussiana es simétrica y nula, la t-Student es simétrica y en la Gúmbel sólo existe dependencia en la cola derecha. Como se puede apreciar, los valores son bastante parecidos en las colas que hay dependencia.

Tabla 14: Coeficiente de dependencia de la cola caso trivariante: Daños materiales - Daños corporales - Hogar

Cópula	Dependencia en la cola izquierda	Dependencia en la cola derecha
Gaussiana	$\rho_1 = \rho_2 = \rho_3 = 0$	$\rho_1 = \rho_2 = \rho_3 = 0$
	$\rho_1 = 0,10721413$	$\rho_1 = 0,10721413$
t-Student	$\rho_2 = 0,09851446$	$\rho_2 = 0,09851446$
	$\rho_3 = 0,07090350$	$\rho_3 = 0,07090350$
Gúmbel	0	0.06439229

*Fuente: Elaboración propia*

Posteriormente, deberemos obtener las funciones de distribución marginal de cada uno de los factores de riesgo a partir de la estructura de dependencia ajustada por la cópula t-Student, puesto que ha sido la seleccionada. Observando los histogramas hemos considerado que las distribuciones que mejor ajustarían nuestros datos serán la Lognormal, Weibull, Exponencial, Gamma y Champernowne<sup>6</sup>.

<sup>6</sup>Ver Bolancé, C., et.al (2012b) y Buch-Larsen, T., et. al (2005a)

Tabla 15: Elección marginales caso trivariante: Daños materiales - Daños corporales - Hogar

<b>Marginal</b>	<b>Riesgo</b>	<b>AIC</b>	<b>BIC</b>
Lognormal	cost_md	3223.341	3232.54
	cost_bi	4262.8	4272
	cost_home	-995.0288	-985.829
Weibull	cost_md	2775.763	2784.963
	cost_bi	3982.914	3992.114
	cost_home	-1075.865	-1066.666
Exponencial	cost_md	2775.343	2779.943
	cost_bi	4292.057	4296.657
	cost_home	-77.66966	-73.06979
Gamma	cost_md	2775.247	2784.446
	cost_bi	4038.819	4048.019
	cost_home	-1059.471	-1050.271
Champernowne	cost_md	2684.689	2698.489
	cost_bi	3996.411	4010.211
	cost_home	-873.4992	-859.6996

*Fuente: Elaboración propia*

Tal y como puede observarse, para el `coste_md` y `cost_bi`, la distribución Champernowne es la que arroja un AIC y BIC menor. Por otro lado, siguiendo el mismo criterio, la distribución Weibull será la más adecuada para `cost_home`.

Por ello, las distribuciones que hemos elegido, serán las que capturan de mejor manera el comportamiento, incorporando el máximo de datos y sobre todo los valores extremos. Una vez definidas las marginales de los tipos de riesgo, construiremos la distribución multivariante a partir del comportamiento conjunto de las variables y la cópula ajustada.

Para cuantificar el riesgo de la pérdida total a partir de las cópulas, aplicamos Monte Carlo simulando 1,000,000 réplicas. En la Tabla 16, se muestran los resultados de las estimaciones del VaR para las diferentes  $\alpha=0.99, 0.995, 0.999$ . Observamos cada valor para las cinco distribuciones elegidas:

Tabla 16: VaR- Cópula t-Student, caso trivariante: Daños materiales - Daños corporales - Hogar

<b>Marginal</b>	<b>VaR<sub>99%</sub></b>	<b>VaR<sub>99,5%</sub></b>	<b>VaR<sub>99,9%</sub></b>
Lognormal	319.8964	525.1502	1435.8841
Weibull	53.13419	64.68022	93.54762
Exponencial	36.57644	41.72947	53.76345
Gamma	3.436904	4.068636	5.641240
Champernowne	259.9302	781.1205	10117.2367
REF	55.01688	74.24688	144.01798

*Fuente: Elaboración propia*

Si aplicamos la cópula t-Student con marginales Lognormal, nos devuelve el valor más elevado del VaR al 99%, pero si incrementamos a un 99.5% o a 99.9% el VaR más alto nos lo presenta la cópula Gúmbel con distribuciones marginales Champernowne.

El caso *REF* se trata de la cuantificación a partir de la unión de la cópula t-Student -siendo la cópula óptima- y las marginales que mejor han ajustado -siendo para el *cost\_md* y *cost\_bi*, la distribución Champernowne y para *coste\_home* la Weibull-. Por ello, la podemos tomar como la referencia puesto que según criterios de selección será la que ajusta de una manera más eficiente.

De tal forma que, la cópula t-Student con la distribución Weibull sería la que más se aproxima al valor de *REF*. En cambio, con la distribución Lognormal o Champernowne el valor está muy por encima.

Aplicar diferentes tipos de distribuciones marginales puede hacer variar la estimación de los niveles de riesgo de la cartera, ya que los resultados difieren para cada tipo de modelo. Por lo tanto, de ahí la importancia de una buena elección del modelo, puesto que se puede subestimar o sobreestimar el riesgo de la cartera.

#### 4.4.2. Caso bivariante: Daños corporales - Hogar

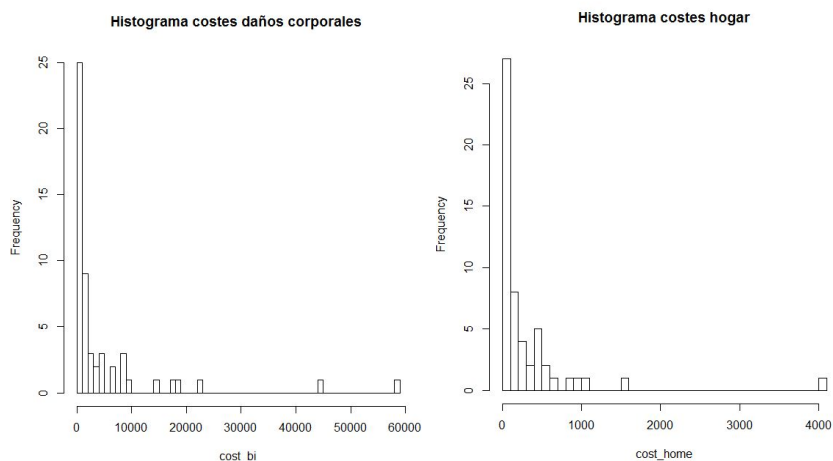
Para este caso, se han seleccionado las observaciones en que se haya declarado siniestro de daños corporales y de hogar. Nos quedamos con 54 observaciones.

Tabla 17: Estadísticos descriptivos caso bivalente: Daños corporales - Hogar

Variabes	Media	Desv.Estándard	Asimetría	Mínimo	Máximo	Mediana
cost_bi	294.7	10507.95	3.4641	0	58056.2	1194.5
cost_home	5067.9	609.5102	4.572431	0	4063.3	111.0

*Fuente: Elaboración propia*

Figura 8: Histogramas: Daños corporales, Hogar



*Fuente: Elaboración propia*

La Tabla 17, describe como cost\_bi incurre en unos costes mucho más elevados que cost\_home, esto también se observa en la dispersión, puesto que, cost\_bi es mucho mayor, reflejándose en los histogramas de la Figura 8. Además, observamos como la asimetría es de un valor similar y positiva.

Tabla 18: Estimación parámetros caso bivalente: Daños corporales - Hogar

Cópuas	Parámetros iniciales	Parámetros finales	Loglikelihood
Gaussian	$\rho = -0,1227156$	$\rho = -0,4124$	3.426
t-Student	$\rho = -0,1227156$ df=4	$\rho = -0,4176$ df=4.7365	3.809

*Fuente: Elaboración propia*

La cópula que presenta una mayor verosimilitud por lo que respecta a sus parámetros es la t-Student, teniendo en cuenta que la Gaussiana devuelve un valor muy cercano.

Tabla 19: Estimación cópula caso bivalente: Daños corporales - Hogar

Cópula	AIC	BIC
Gaussiana	-4.85293	-2.863946
t-Student	-3.617616	0.3603524
Gúmbel	2.000001	3.988985

*Fuente: Elaboración propia*

La cópula seleccionada por tener un menor AIC y BIC es la Gaussiana.

Tabla 20: Coeficiente de dependencia de la cola caso bivalente: Daños corporales - Hogar

Cópula	Dependencia en la cola izquierda	Dependencia en la cola derecha
Gaussiana	0	0
t-Student	0.0525564	0.0525564

*Fuente: Elaboración propia*

En la Tabla 20, observamos que el coeficiente de dependencia de la cola para la Gaussiana es nulo, para la t-Student es el mismo en la cola derecha que en la cola izquierda por su simetría y para la Gúmbel no se ha podido ajustar.



Tabla 21: Elección marginales caso bivalente: Daños corporales - Hogar

<b>Marginal</b>	<b>Riesgo</b>	<b>AIC</b>	<b>BIC</b>
Lognormal	cost_bi	240.6278	244.6057
	cost_home	-92.52843	-88.55046
Weibull	cost_bi	227.6925	231.6705
	cost_home	-97.0629	-93.08493
Exponencial	cost_bi	285.2966	287.2856
	cost_home	-21.58109	-19.59211
Gamma	cost_bi	230.3473	234.3252
	cost_home	-97.81708	-93.83912
Champernowne	cost_bi	237.0507	243.0177
	cost_home	-79.30628	-73.33933

*Fuente: Elaboración propia*

Tal y como se evidencia en la Tabla 21, la distribución marginal que mejor ajusta los costes de daños corporales es la Weibull, para los costes de hogar, es la Champernowne.

En la siguiente tabla, se muestra el VaR estimado a diferentes niveles de confianza:

Tabla 22: VaR- Cópula Gaussiana, caso bivalente: Daños corporales - Hogar

<b>Marginal</b>	<b>VaR<sub>99%</sub></b>	<b>VaR<sub>99,5%</sub></b>	<b>VaR<sub>99,9%</sub></b>
Lognormal	376.6089	722.4445	2797.4362
Weibull	52.54260	69.49572	118.36499
Exponencial	23.39565	26.95886	35.26261
Gamma	2.982867	3.616989	5.158219
Champernowne	1707.779	5115.567	71339.373
REF	482.3332	1697.0652	32385.2855

*Fuente: Elaboración propia*

Considerando los resultados, la cuantificación del riesgo con la cópula Gaussiana y distribuciones marginales tanto Gamma, como Weibull o Exponencial nos devuelven unos VaR<sub>99%</sub>, VaR<sub>99,5%</sub> y VaR<sub>99,9%</sub> con unos valores pequeños en comparación a aplicar la distribución marginal Lognormal y sobre todo Champernowne. Si nos fijamos en los re-

sultados de *REF*, podemos ver como todos los modelos infravaloran el riesgo excepto con el caso de la Champernowne que sucede lo contrario.

#### 4.4.3. Caso bivariante: Daños materiales - Daños corporales

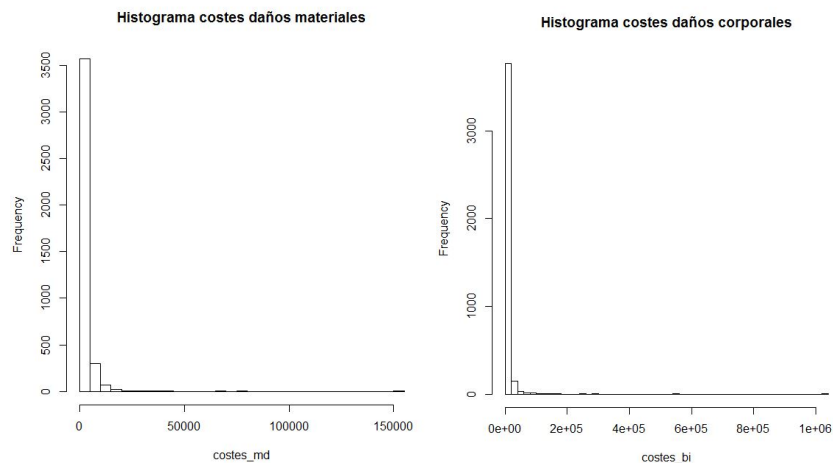
En este caso, se han seleccionado los individuos que han declarado siniestros de automóvil tanto en daños materiales como en daños corporales. Nos hemos quedado con 3981 observaciones.

Tabla 23: Estadísticos descriptivos caso bivariante: Daños materiales - Daños corporales

Variabes	Media	Desv.Estándard	Asimetría	Mínimo	Máximo	Mediana
cost_md	2496	4039.2	17.04408	0	153178	1676
cost_bi	6955	23203.57	26.60952	0	1023100	3312

*Fuente: Elaboración propia*

Figura 9: Histogramas: Daños materiales, Daños corporales



*Fuente: Elaboración propia*

En vista a los resultados obtenidos en la Tabla 23 y la Figura 9, determinamos que en ambas variables hay asimetría fuerte a la derecha. Entre el valor mínimo y máximo

hay mucha diferencia, provocando una gran dispersión de los datos, tal y como podemos observar en los histogramas.

Tabla 24: Estimación parámetros caso bivalente: Daños materiales - Daños corporales

<b>Cóputas</b>	<b>Parámetros iniciales</b>	<b>Parámetros finales</b>	<b>Loglikelihood</b>
Gaussian	$\rho = 0,03710382$	$\rho = 0,1304$	31.67
t-Student	$\rho = 0,03710382$ df=4	$\rho = 0,1296$ df=9.1092	45.52
Gúmbel	1.09006	1.08	34.38

*Fuente: Elaboración propia*

El valor del logaritmo de máxima-verosimilitud mayor corresponde a los parámetros de la cóputa de t-Student.

Tabla 25: Elección cóputa caso bivalente: Daños materiales - Daños corporales

<b>Cóputa</b>	<b>AIC</b>	<b>BIC</b>
Gaussiana	-61.34977	-55.06048
t-Student	-87.03195	-74.45337
Gúmbel	-66.75192	-60.46263

*Fuente: Elaboración propia*

Para esta combinación, seleccionamos la cóputa t-Student.

Tabla 26: Coeficiente de dependencia de la cola caso bivalente: Daños materiales - Daños corporales

<b>Cóputa</b>	<b>Dependencia en la cola izquierda</b>	<b>Dependencia en la cola derecha</b>
Gaussiana	0	0
t-Student	0.08375816	0.08375816
Gúmbel	0	0.1113165

*Fuente: Elaboración propia*

Como en los casos anteriores comprobamos que la dependencia de la Gaussiana y la t-Student es simétrica, siendo para la primera nula. La cóputa Gúmbel sólo tiene depen-

dencia en la cola derecha. De nuevo, los valores de la dependencia en la cola derecha de la Gúmbel y la t-Student son muy parecidos.

Tabla 27: Elección marginales caso bivalente: Daños materiales - Daños corporales

<b>Marginal</b>	<b>Riesgo</b>	<b>AIC</b>	<b>BIC</b>
Lognormal	cost_md	17382.6	17395.17
	cost_bi	23371.67	23384.25
Weibull	cost_md	15221.98	15234.56
	cost_bi	21592.6	21580.02
Exponencial	cost_md	15250.49	15256.78
	cost_bi	23406.9	23413.19
Gamma	cost_md	15241.3	15253.88
	cost_bi	21758.67	21771.25
Champernowne	cost_md	14506.47	14525.33
	cost_bi	21928.8	21947.67

*Fuente: Elaboración propia*

De acuerdo con la Tabla 27, la distribución marginal que mejor ajusta los datos de los costes de daños corporales es la distribución Weibull y para los costes de daños materiales es la distribución Champernowne.

Tabla 28: VaR- Cópula t-Student, caso bivalente: Daños materiales - Daños corporales

<b>Marginal</b>	<b>VaR<sub>99%</sub></b>	<b>VaR<sub>99,5%</sub></b>	<b>VaR<sub>99,9%</sub></b>
Lognormal	425.3562	730.9219	2311.3220
Weibull	55.67216	68.19757	101.45444
Exponencial	36.40719	41.52337	53.85289
Gamma	1.738816	2.003296	2.615425
Champernowne	59.26205	81.52938	170.38278
REF	56.10053	68.80124	103.81200

*Fuente: Elaboración propia*

Si tomamos el valor *REF*, las distribuciones marginales, con cópula t-Student, que más se acercan al resultado obtenido como referencia son la distribución Weibull y la Champernowne, aunque para el VaR<sub>99,5%</sub> la cópula con distribuciones marginales Weibull sigue

siendo muy similar, en cambio la Champernowne deja de serlo. Cabe destacar que si ajustamos con distribución marginal Gamma estaríamos infravalorando el riesgo y por el contrario, con marginales Lognormal lo estaríamos sobrevalorando.

#### 4.4.4. Caso bivalente: Daños materiales - Hogar

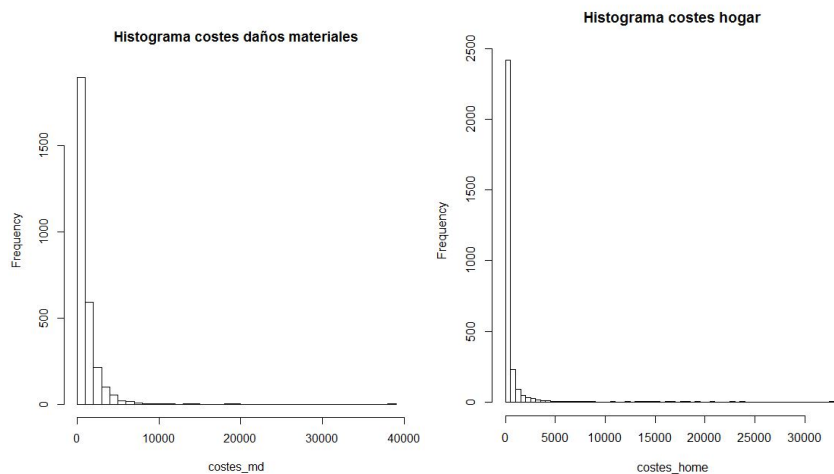
Este caso, tomamos las variables daños materiales y hogar. Las observaciones con las que trabajamos son 2905.

Tabla 29: Estadísticos descriptivos caso bivalente: Daños materiales - Hogar

VARIABLES	Media	Desv.Estándar	Asimetría	Mínimo	Máximo	Mediana
cost_md	1375	1435.742	9.239151	0	38758	882
cost_home	468.0	1753.493	10.55789	0	33284.6	124.4

*Fuente: Elaboración propia*

Figura 10: Histogramas: Daños materiales, Hogar



*Fuente: Elaboración propia*

Observamos que tanto los estadísticos descriptivos de la Tabla 29, como los histogramas de la Figura 10, son bastante similares por lo que respecta a los dos factores de riesgo.

Presentan una asimetría muy parecida, al igual que el valor máximo y por consiguiente la dispersión.

Tabla 30: Estimación parámetros caso bivalente: Daños materiales - Hogar

Cóputas	Parámetros iniciales	Parámetros finales	Loglikelihood
Gaussian	$\rho = 0,00358668$	$\rho = 5,388e - 05$	3.224e-06
t-Student	$\rho = 0,00358668$ df=4	$\rho = -0,0007346$ df=79.6031743	-0.03106
Gúmbel	-	1.003	0.08885

*Fuente: Elaboración propia*

El parámetro de la cóputa Gúmbel es el que presenta una mayor verosimilitud.

Tabla 31: Elección cóputa caso bivalente: Daños materiales - Hogar

Cóputa	AIC	BIC
Gaussiana	1.999994	7.974182
t-Student	4.06212	16.0105
Gúmbel	1.822303	7.796491

*Fuente: Elaboración propia*

La cóputa que mejor ajusta en este caso será la Gúmbel.

Tabla 32: Coeficiente de dependencia de la cola caso bivalente: Daños materiales - Hogar

Cóputa	Dependencia en la cola izquierda	Dependencia en la cola derecha
Gaussiana	0	0
t-Student	0.07635067	0.07635067

*Fuente: Elaboración propia*

Otra vez se vuelve a poner de manifiesto que el coeficiente de dependencia para la cóputa Gaussiana es nulo, y para la t-Student tenemos el mismo valor en la cola superior que en la inferior y para la Gúmbel tampoco se ha podido ajustar.

Tabla 33: Elección marginales caso bivalente: Daños materiales - Hogar

<b>Marginal</b>	<b>Riesgo</b>	<b>AIC</b>	<b>BIC</b>
Lognormal	cost_md	9631.035	9642.983
	cost_home	-4135.74	-4123.791
Weibull	cost_md	7531.962	7543.911
	cost_home	-4287.118	-4275.169
Exponencial	cost_md	7664.895	7670.869
	cost_home	1412.938	1418.912
Gamma	cost_md	7541.481	7553.429
	cost_home	-3935.809	-3923.861
Champernowne	cost_md	6746.483	6764.405
	cost_home	-3625.999	-3608.076

*Fuente: Elaboración propia*

Las distribuciones marginales que mejor ajustan los costes por siniestros de hogar y los costes de daños materiales son la distribución Weibull y la Champernowne, respectivamente.

Tabla 34: VaR- Cópula Gumbel, caso bivalente: Daños materiales - Hogar

<b>Marginal</b>	<b>VaR<sub>99%</sub></b>	<b>VaR<sub>99,5%</sub></b>	<b>VaR<sub>99,9%</sub></b>
Lognormal	43.17871	71.1649	230.89419
Weibull	7.78805	9.73544	15.93588
Exponencial	6.926278	7.884815	10.151725
Gamma	5.346163	6.046839	7.704946
Champernowne	339.544	1136.013	16262.066
REF	7.491062	9.667056	16.202547

*Fuente: Elaboración propia*

Los resultados reflejan que la cópula Gumbel con distribuciones marginales Gamma o Weibull o Exponencial serán las que mejor cuantificarán el riesgo, así pues arrojan valores muy parecidos a los de *REF* para el  $VaR_{99\%}$ . Por lo que respecta a los otros niveles de significación más cercanos al cuantil 1, es la aplicación de la distribución Weibull la que sigue siendo más similar. Todo lo contrario ocurre cuando se usan las distribuciones Lognormal o Champernowne, mostrándonos valores muy por encima de *REF*, en consecuencia sobreestimarían el riesgo.

#### 4.4.5. Caso univariante: Daños materiales

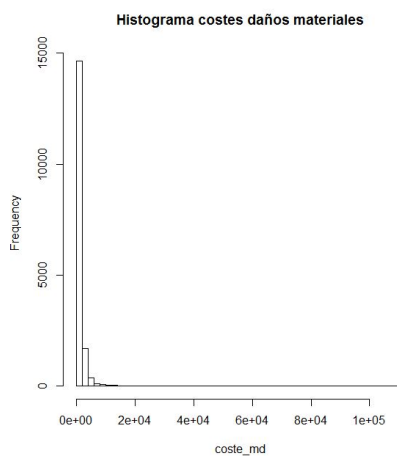
Nos quedamos con datos en los que sólo haya habido coste de daños materiales y son 16926 observaciones.

Tabla 35: Estadísticos descriptivos caso univariante: Daños materiales

Variables	Media	Desv.Estándar	Asimetría	Mínimo	Máximo	Mediana
coste_md	1346	1609.94	23.33188	0	111400	882

*Fuente: Elaboración propia*

Figura 11: Histograma Daños Materiales



*Fuente: Elaboración propia*

Analizando el resultado tanto del histograma como de la Tabla 35, vemos como los datos presentan una asimetría muy marcada a la derecha.



Tabla 36: Elección marginales caso univariante: Daños materiales

<b>Marginal</b>	<b>AIC</b>	<b>BIC</b>
Lognormal	56160.64	56176.11
Weibull	43495.8	43511.28
Exponencial	43928.25	43935.99
Gamma	43570.19	43585.66
Champernowne	38295.41	38318.62

*Fuente: Elaboración propia*

Según la Tabla 36, la distribución marginal que mejor ajusta los datos es la Champernowne.

Tabla 37: VaR, caso univariante: Daños materiales

<b>Marginal</b>	<b>VaR<sub>99%</sub></b>	<b>VaR<sub>99,5%</sub></b>	<b>VaR<sub>99,9%</sub></b>
Lognormal	27.75008	40.83863	90.25955
Weibull	5.452020	6.195709	7.837931
Exponencial	6.206409	7.162154	9.253710
Gamma	5.646616	6.454607	8.289985
Champernowne	4.316867	5.340703	8.698402

*Fuente: Elaboración propia*

Para el cálculo del VaR univariante simplemente hemos ajustado los datos a las distribuciones que se están aplicando. Observamos que excepto la distribución Lognormal que presenta un VaR mucho más elevado, el resto devuelven resultados muy parecidos.

#### 4.4.6. Caso univariante: Daños corporales

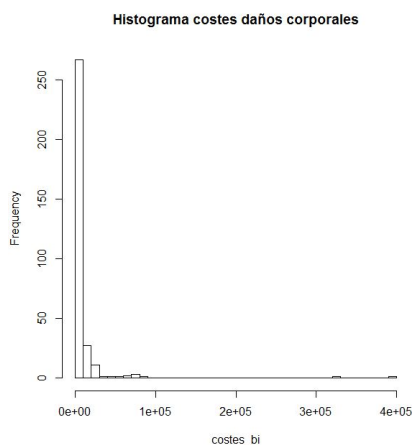
Para este caso tenemos 316 observaciones, puesto que nos quedamos con los individuos que sólo han tenido siniestro de daños corporales

Tabla 38: Estadísticos descriptivos caso univariante: Daños corporales

Variabes	Media	Desv.Estándard	Asimetría	Mínimo	Máximo	Mediana
cost_md	7801.5	30367.33	10.09552	0	392352.2	1123.4

*Fuente: Elaboración propia*

Figura 12: Histograma Daños corporales



*Fuente: Elaboración propia*

De nuevo, el resultado indica que los datos son asimétricos hacia la derecha. En comparación con los datos univariantes de daños materiales, estos datos muestran menor asimetría. Por lo que respecta a la desviación estándar en este caso es bastante más elevada.

Tabla 39: Elección marginales caso univariante: Daños corporales

Marginal	AIC	BIC
Lognormal	1433.357	1440.869
Weibull	1366.733	1374.245
Exponencial	1932.408	1936.163
Gamma	1408.154	1415.665
Champernowne	1414.984	1426.251

*Fuente: Elaboración propia*

Según los criterios de información la distribución marginal que debemos elegir es la Weibull.

Tabla 40: VaR, caso univariante: Daños corporales

<b>Marginal</b>	<b>VaR<sub>99%</sub></b>	<b>VaR<sub>99,5%</sub></b>	<b>VaR<sub>99,9%</sub></b>
Lognormal	662.325	1359.771	5761.798
Weibull	84.11704	114.70248	203.95743
Exponencial	35.64215	40.99784	53.27999
Gamma	68.11243	83.14189	118.23924
Champernowne	940.5661	2610.5679	28728.6567

*Fuente: Elaboración propia*

En la Tabla 40, se observan resultados bastante distintos para cada distribución. Si tomamos como referencia la Weibull, por ser la más eficiente, vemos como la distribución Lognormal y Champernowne devuelven un valor muy por encima. En cambio con la distribución Exponencial o con la Gamma estaríamos subestimando el riesgo.

#### 4.4.7. Caso univariante: Hogar

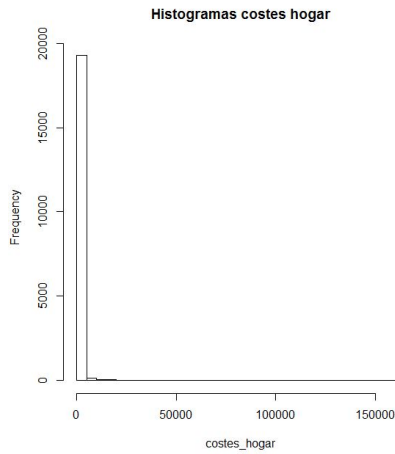
En este último caso, tenemos los datos en los que el individuo ha declarado sólo en el hogar. Hay 19513 observaciones.

Tabla 41: Estadísticos descriptivos caso univariante: Hogar

<b>Variabes</b>	<b>Media</b>	<b>Desv.Estándard</b>	<b>Asimetría</b>	<b>Mínimo</b>	<b>Máximo</b>	<b>Mediana</b>
cost_home	442.4	2573.053	33.32864	0	161700.0	118.8

*Fuente: Elaboración propia*

Figura 13: Histograma Hogar



*Fuente: Elaboración propia*

De los tres casos univariantes, los costes de hogar son los que reflejan una mayor asimetría positiva.

Tabla 42: Elección marginales caso univariante: Hogar

<b>Marginal</b>	<b>AIC</b>	<b>BIC</b>
Lognormal	-30511.69	-30495.94
Weibull	-31538.03	-31522.27
Exponencial	7287.789	7295.668
Gamma	-28715.15	-28699.39
Champernowne	-26978.95	-26955.32

*Fuente: Elaboración propia*

De nuevo el modelo Weibull es el mejor para ajustar los datos.

Tabla 43: VaR, caso univariante: Hogar

<b>Marginal</b>	<b>VaR<sub>99%</sub></b>	<b>VaR<sub>99,5%</sub></b>	<b>VaR<sub>99,9%</sub></b>
Lognormal	23.56559	47.68161	202.56170
Weibull	5.018669	6.987516	12.841036
Exponencial	2.042854	2.357543	3.046210
Gamma	3.985844	4.862384	6.990720
Champernowne	323.1366	1071.1748	17891.3311

*Fuente: Elaboración propia*

Las reflexiones del VaR ajustando con distribución marginal Weibull, son las mismas que en el anterior caso. Vemos que la distribución Champernowne estaría sobreestimando de forma exagerada el riesgo, seguida por la Lognormal. Estos valores se agravan a medida que aumentamos el nivel de confianza. En el caso de la Exponencial y Gamma sucede lo contrario pero de forma más moderada.

Puntualizar que para cada combinación se ha calculado el TVaR para los diferentes niveles de confianza  $\alpha=0.99, 0.995, 0.999$  y como se han extraído las mismas conclusiones que con los respectivos VaR se han adjuntado en el Anexo I.

## 4.5. Riesgo de pérdida integrado

Podemos establecer que el comportamiento del riesgo de pérdida integrado, compuesto por dos variables estocásticas. Por un lado  $L$  que toma valores no-negativos y la variable binaria  $y$  que equivale a 1 si el tomador declara el siniestro y 0 en caso contrario.

Para la obtención del riesgo de pérdida de cada perfil se ha optado por aplicar el *Value at Risk* correspondiente a cada combinación (sección 4.4) a un nivel de confianza  $\alpha=0.995$ , ya que como se ha comentado anteriormente es el nivel exigido para las aseguradoras por la normativa vigente. Cabe recordar que los resultados de los VaR se muestran en miles y en consecuencia también la  $L$ . Por su parte, la probabilidad asociada a cada combinación será la obtenida en el apartado 4.3.

$$\begin{aligned}
L = & VaR_L(C_{md}, C_{bi}, C_{home}) \times Pr(y_{md} = 1, y_{bi} = 1, y_{home} = 1) + \\
& VaR_L(C_{md}, C_{bi}) \times Pr(y_{md} = 1, y_{bi} = 1, y_{home} = 0) + \\
& VaR_L(C_{md}, C_{home}) \times Pr(y_{md} = 1, y_{bi} = 0, y_{home} = 1) + \\
& VaR_L(C_{bi}, C_{home}) \times Pr(y_{md} = 0, y_{bi} = 1, y_{home} = 1) + \\
& VaR_L(C_{md}) \times Pr(y_{md} = 1, y_{bi} = 0, y_{home} = 0) + \\
& VaR_L(C_{bi}) \times Pr(y_{md} = 0, y_{bi} = 1, y_{home} = 0) + \\
& VaR_L(C_{home}) \times Pr(y_{md} = 0, y_{bi} = 0, y_{home} = 1)
\end{aligned}$$

PERFIL 1:

$$\begin{aligned}
L = & 74,24688 \times 0,4115549 + 68,80124 \times 0,05771433 + 9,667056 \times 0,2839457 + \\
& 1697,0652 \times 0,00782554 + 5,340703 \times 0,04510453 + 114,70248 \times 0,001220068 + \\
& 6,987516 \times 0,1639832 = 52,07952529
\end{aligned}$$

PERFIL 2:

$$\begin{aligned}
L = & 74,24688 \times 0,4252129 + 68,80124 \times 0,06779071 + 9,667056 \times 0,2737775 + \\
& 1697,0652 \times 0,007295146 + 5,340703 \times 0,04942243 + 114,70248 \times 0,001293901 + \\
& 6,987516 \times 0,1462587 = 52,69612608
\end{aligned}$$

PERFIL 3:

$$\begin{aligned}
L = & 74,24688 \times 0,3633445 + 68,80124 \times 0,05106907 + 9,667056 \times 0,3402907 + \\
& 1697,0652 \times 0,004047292 + 5,340703 \times 0,05431758 + 114,70248 \times 0,0006334073 + \\
& 6,987516 \times 0,1583129 = 42,11789986
\end{aligned}$$

PERFIL 4:

$$\begin{aligned}
L = & 74,24688 \times 0,3092828 + 68,80124 \times 0,03877884 + 9,667056 \times 0,3104827 + \\
& 1697,0652 \times 0,008417038 + 5,340703 \times 0,04390415 + 114,70248 \times 0,00116299 + \\
& 6,987516 \times 0,2490094 = 45,02486517
\end{aligned}$$

Tabla 44: Resultados riesgo de pérdida con  $\alpha=0.995$  integrado para los perfiles

<b>PERFIL 1</b>	<b>PERFIL 2</b>	<b>PERFIL 3</b>	<b>PERFIL 4</b>
52.07952529	52.69612608	42.11789986	45.02486517

*Fuente: Elaboración propia*

Podemos ver que para los perfiles 3 y 4 el riesgo de pérdida es menor respecto del 1 y 2. Esto refleja que el factor de ser hombre genera un menor riesgo en nuestro modelo, así como para el caso del perfil 3 vivir en el norte de España y para el caso del perfil 4 la antigüedad del permiso de conducir (41 años). Por contra los perfiles 1 y 2 se ven penalizados por el hecho de ser mujeres, y para el perfil 2 además se suma la poca experiencia en la conducción (1 año de antigüedad de carnet). Estos resultados concuerdan con la información de la Tabla 8.

## 5. Conclusiones

Después de realizar este análisis podemos extraer que es importante tener en cuenta que los clientes que contratan distintas pólizas en distintas líneas de negocio tienen un comportamiento que nos indica que el riesgo en estas líneas de negocio está relacionado y por ello la aplicación del modelo multivariante será la mejor opción. Para la modelización del hecho de declarar siniestro se ha utilizado el modelo Probit Multivariante.

Para modelizar los costes, hemos utilizado las cópulas. En este estudio, las cópulas que mejor han ajustado han sido la t-Student y la Gúmbel ambas son de valor extremo. El uso de modelos de cópulas nos da una gran versatilidad a la hora de comparar diferentes estructuras de dependencia entre las tres líneas de negocio. También en la valoración de la incidencia de cada una de ellas en la distribución de la pérdida y en consecuencia, de la medida de riesgo.

En este caso, las distribuciones marginales de los factores de riesgo tienen asimetría a la derecha y hemos comprobado que las distribuciones Champernowne y Weibull son las que recogen mejor todos los datos de la distribución.

Destacamos la importancia que supone ajustar de manera individual las distribuciones de las marginales que conformarán las cópulas. De forma que no todas las distribuciones tienen porque seguir el mismo modelo y por tanto ajustar la cópula con las marginales correctas evitará que se sobrevalore o infravalore el riesgo de pérdida. Es decir, según la selección puede hacer variar mucho los resultados. Así pues, el valor de la medida de riesgo asociado a la cartera, como se ha observado en los resultados, presenta diferencias significativas según la estructura de dependencia ajustada a las líneas de negocio.

Hemos de tener en cuenta que la principal misión de las aseguradoras es gestionar el riesgo de pérdida. Para ello, es necesaria una correcta valoración que nos permita poder explicar y predecir el comportamiento de la cartera que transfieren los asegurados, sobre todo en la severidad de los accidentes no-vida. Ésto garantizará la estabilidad y solidez económica de la compañía.



# Anexo

## I. Tablas TVaR

Tabla 45: TVaR- Cópula t-Student, caso trivariante: Daños materiales - Daños corporales - Hogar

<b>Marginal</b>	<b>TVaR<sub>99%</sub></b>	<b>TVaR<sub>99,5%</sub></b>	<b>TVaR<sub>99,9%</sub></b>
Lognormal	847.4997	1291.792	3210.497
Weibull	70.82589	83.48656	115.4823
Exponencial	44.08406	49.31538	61.53805
Gamma	4.367047	5.017504	6.583453
Champernowne	65658.55	130878.4	643941.1

*Fuente: Elaboración propia*

Tabla 46: TVaR- Cópula Gaussiana, caso bivalente: Daños corporales - Hogar

<b>Marginal</b>	<b>TVaR<sub>99%</sub></b>	<b>TVaR<sub>99,5%</sub></b>	<b>TVaR<sub>99,9%</sub></b>
Lognormal	1628.204	2743.828	8529.842
Weibull	80.32618	100.8042	158.4277
Exponencial	28.54883	32.12486	40.59233
Gamma	3.909668	4.55955	6.097725
Champernowne	8192853	16382799	81842420

*Fuente: Elaboración propia*

Tabla 47: TVaR- Cópula t-Student, caso bivalente: Daños materiales - Daños corporales

<b>Marginal</b>	<b>TVaR<sub>99%</sub></b>	<b>TVaR<sub>99,5%</sub></b>	<b>TVaR<sub>99,9%</sub></b>
Lognormal	1337.908	2125.039	5847.204
Weibull	75.08984	88.92582	125.3516
Exponencial	43.94757	49.22833	62.15822
Gamma	2.120732	2.387093	3.014651
Champernowne	109.3141	149.9637	313.5906

*Fuente: Elaboración propia*

Tabla 48: TVaR- Cópula Gúmbel, caso bivalente: Daños materiales - Hogar

<b>Marginal</b>	<b>TVaR<sub>99%</sub></b>	<b>TVaR<sub>99,5%</sub></b>	<b>TVaR<sub>99,9%</sub></b>
Lognormal	139.2639	223.987	653.2754
Weibull	11.16353	13.73652	21.3316
Exponencial	8.325629	9.308115	11.60895
Gamma	6.372982	7.090452	8.768153
Champernowne	157670	314725.9	1557052

*Fuente: Elaboración propia*

Tabla 49: TVaR, caso univariante: Daños materiales

<b>Marginal</b>	<b>TVaR<sub>99%</sub></b>	<b>TVaR<sub>99,5%</sub></b>	<b>TVaR<sub>99,9%</sub></b>
Lognormal	53.89545	74.51785	146.2896
Weibull	6.492705	7.204708	8.770769
Exponencial	7.544761	8.466755	10.52602
Gamma	6.802772	7.594458	9.443752
Champernowne	6.190983	7.63226	12.34093

*Fuente: Elaboración propia*

Tabla 50: TVaR, caso univariante: Daños corporales

<b>Marginal</b>	<b>TVaR<sub>99%</sub></b>	<b>TVaR<sub>99,5%</sub></b>	<b>TVaR<sub>99,9%</sub></b>
Lognormal	3603.749	6274.017	21088.38
Weibull	135.5128	173.8532	284.958
Exponencial	43.36041	48.71219	61.59074
Gamma	90.15964	105.6673	143.4693
Champernowne	3107595	6213652	31036465

*Fuente: Elaboración propia*

Tabla 51: TVaR, caso univariante: Hogar

<b>Marginal</b>	<b>TVaR<sub>99%</sub></b>	<b>TVaR<sub>99,5%</sub></b>	<b>TVaR<sub>99,9%</sub></b>
Lognormal	113.4991	193.8835	610.6098
Weibull	8.306807	10.75734	17.94196
Exponencial	2.483521	2.787096	3.46513
Gamma	5.288699	6.203849	8.370072
Champernowne	20522657	41044734	205207036

*Fuente: Elaboración propia*

## II. Código SAS: Aplicación Probit Multivariante

```

libname tfm '.';

data tfm.datos2;
set tfm.datos1;
if nclaims_md>0 then nclaims_mdd=1; else nclaims_mdd=0;
if nclaims_bi>0 then nclaims_bid=1; else nclaims_bid=0;
if nclaims_home>0 then nclaims_homed=1; else nclaims_homed=0;
if client_sex='Man' then client_sex=1; else if client_sex='Woman' then client_sex=0;
client_sex2 = input(client_sex, best12.);
client_age100=client_age/100;
lic_age100=lic_age/100;
if (north = 1) then zona='north';
if (rest = 1) then zona='rest';
if (cities2= 1) then zona='cities2';
run;

data tfm.datos;
set tfm.datos2;
pexposiauto=(exposi_auto/115926.00)*0.9;
pexposihome=(exposi_home/226164.00)*0.9;
expoauto=probit(pexposiauto);
expohome=probit(pexposihome);
run;

proc qlim data=tfm.datos method=qn;
class zona;
model nclaims_mdd = expoauto client_sex2 client_age100 lic_age100
client_nother zona;

model nclaims_bid = expoauto client_sex2 client_age100 lic_age100
client_nother zona;

model nclaims_homed = expohome client_sex2 client_age100
client_nother zona;

endogenous nclaims_mdd nclaims_bid nclaims_homed ~ discrete ;
restrict nclaims_mdd.expoauto=1, nclaims_bid.expoauto=1, nclaims_homed.expohome=1;
run;

```

### III. Código R: Aplicación cópulas

```
#####
# Instalación de los packages

install.packages("copula")
library(copula)

install.packages("fCopulae")
library(fCopulae)

install.packages("QRM")
library(QRM)

install.packages("MASS")
library(MASS)

##### CASO TRIVARIANTE: MD-BI-HOME #####
data<-read.table("Datos TFM.csv", header=TRUE, sep=";")
head(data)
costes<-cbind(data[9],data[11],data[15])
head (costes)
attach(costes)
cost<-as.matrix(costes)
cost123<-cost[(data$nclaims_md>0&data$nclaims_bi>0&data$nclaims_home>0),]
head(cost123)
nrow(cost123)
summary(cost123)

#Desviación estándar
sd(cost123[,1])
sd(cost123[,2])
sd(cost123[,3])

#Asimetría
skewness(cost123[,1])
skewness(cost123[,2])
skewness(cost123[,3])

#Histograma
hist(cost123[,1],50, xlab= "Coste_md",main="Histograma costes daños materiales")
hist(cost123[,2],50, xlab= "Coste_home",main="Histograma costes hogar")
hist(cost123[,3],50, xlab= "Coste_bi",main="Histograma costes daños corporales")

cost123<-(cost123/1000)+0.001 # Añadimos 0.001 para que nos coja bien en las marginales.
cost123md<-cost123[,1]
cost123home<-cost123[,2]
cost123bi<-cost123[,3]
n<-nrow(cost123);n

###Transformación de las variables originales a variables uniformes usando la función de distribución empírica###

Ucost123<-apply(cost123,2,edf,adjust=1)
head(Ucost123)
Ucost123_2<-pobs(cost123)
head(Ucost123_2)
```

```

###Cálculo de las correlaciones estimadas por el método clásico (correlación lineal)###
matrizpearson<-cor(cost123)
matrizpearson

###Estimar los coeficientes de correlación Tau y Spearman, entre los tres costes
rotau<-Kendall(cost123)
rotau

###1.Definición parámetros iniciales###

#Parámetros iniciales-Cópula Gumbel
#Método momentos#
ParGum1<-1/(1-rotau[1,2])
ParGum1
ParGum2<-1/(1-rotau[1,3])
ParGum2
ParGum3<-1/(1-rotau[2,3])
ParGum3
#ParGum<-ParGum1 #cojo el valor medio, ya que con el mínimo no está dentro del rango de valores posibles
ParGum<-ParGum1

p1<-matrizpearson[1,2]
p2<-matrizpearson[1,3]
p3<-matrizpearson[2,3]

#2.Definición el objeto cópula

# AJUSTE CÓPULA Y ESTIMACIÓN DE LOS PARÁMETROS DE CÓPULA #

# Cópula Gauss#
norm.cop <- normalCopula(c(p1, p2, p3), dim = 3, dispstr = "un")
norm.cop
NormCopEst<-fitCopula(norm.cop,Ucost123, method="mpl")
NormCopEst

# Cópula t-Student ##### Cópulas de valor extremo #
tCop <- tCopula(c(p1, p2, p3), dim = 3, dispstr="un", df.fixed=FALSE)
tCop
TCopEst<-fitCopula(tCop, Ucost123, method="mpl", estimate.variance=TRUE)
TCopEst

# Cópula Gumbel ##### Cópulas de valor extremo #
gumb.cop <- gumbelCopula(NA_real_, dim =3)
gumb.cop
gumbCopEst<-fitCopula(gumb.cop,Ucost123, method="mpl")
gumbCopEst

AIC(NormCopEst)
AIC(TCopEst)
AIC(gumbCopEst)

BIC(NormCopEst)
BIC(TCopEst)
BIC(gumbCopEst)

#dependencia en colas:
TailDepN<-tailIndex(norm.cop)

```

```

TailDepN

TailDept<-tailIndex(tCop)
TailDept

TailDepG<-tailIndex(gumbelCopula(ParGum, dim =3))
TailDepG

# mejor ajuste t-Student

alfa<-c(0.99, 0.995,0.999)
rep=1000000

#####T-STUDENT#####
#####

t.cop_est<-tCopula(c(0.13074732, 0.18843,-0.01045), dim=3,dispstr = "un", df=9.546921)
set.seed(123)
sim.data.tcopula<-rCopula(rep,t.cop_est)

#.Calcular el VaR para la pérdida L a partir de la
#cópula suponiendo marginales LOG-NORMAL con fitdistr:

mod_cost123md_log<-fitdistr(cost123md,"lognormal");
mod_cost123md_log
meanlog_123md<-coefficients(mod_cost123md_log)["meanlog"]
sdlog_123md<-coefficients(mod_cost123md_log)["sdlog"]

AIC(mod_cost123md_log)
BIC(mod_cost123md_log)

mod_cost123bi_log<-fitdistr(cost123bi,"lognormal");
mod_cost123bi_log
meanlog_123bi<-coefficients(mod_cost123bi_log)["meanlog"]
sdlog_123bi<-coefficients(mod_cost123bi_log)["sdlog"]

AIC(mod_cost123bi_log)
BIC(mod_cost123bi_log)

mod_cost123home_log<-fitdistr(cost123home,"lognormal");
mod_cost123home_log
meanlog_123home<-coefficients(mod_cost123home_log)["meanlog"]
sdlog_123home<-coefficients(mod_cost123home_log)["sdlog"]

AIC(mod_cost123home_log)
BIC(mod_cost123home_log)

sim.t.ln.cost123_md <- qnorm(sim.data.tcopula[,1], mean=meanlog_123md, sd=sdlog_123md)
sim.t.ln.cost123_bi <- qnorm(sim.data.tcopula[,3], mean=meanlog_123bi, sd=sdlog_123bi)
sim.t.ln.cost123_home <- qnorm(sim.data.tcopula[,2], mean=meanlog_123home, sd=sdlog_123home)

sim.t.cost123_md_log=exp(sim.t.ln.cost123_md)
sim.t.cost123_bi_log=exp(sim.t.ln.cost123_bi)
sim.t.cost123_home_log=exp(sim.t.ln.cost123_home)

sim.tcop.L_log=(sim.t.cost123_md_log+sim.t.cost123_home_log+sim.t.cost123_bi_log)-0.001

```

```

VaRtCopulaLN<-quantile(sim.tcop.L_log, alfa)
VaRtCopulaLN

#TVaR
mean(sim.tcop.L_log[sim.tcop.L_log > quantile(sim.tcop.L_log, alfa[1])])
mean(sim.tcop.L_log[sim.tcop.L_log > quantile(sim.tcop.L_log, alfa[2])])
mean(sim.tcop.L_log[sim.tcop.L_log > quantile(sim.tcop.L_log, alfa[3])])
cbind(mean(sim.tcop.L_log[sim.tcop.L_log > quantile(sim.tcop.L_log, alfa[1])]),
mean(sim.tcop.L_log[sim.tcop.L_log > quantile(sim.tcop.L_log, alfa[2])]),
mean(sim.tcop.L_log[sim.tcop.L_log > quantile(sim.tcop.L_log, alfa[3])]))

#.Calcular el VaR para la pérdida L a partir de la
#cópula suponiendo marginales WEIBULL con fitdistr:

mod_cost123md_weibull<-fitdistr(cost123md,"weibull");
shapeweibull_123md<-coefficients(mod_cost123md_weibull)["shape"]
scaleweibull_123md<-coefficients(mod_cost123md_weibull)["scale"]

AIC(mod_cost123md_weibull)
BIC(mod_cost123md_weibull)

mod_cost123bi_weibull<-fitdistr(cost123bi,"weibull");
shapeweibull_123bi<-coefficients(mod_cost123bi_weibull)["shape"]
scaleweibull_123bi<-coefficients(mod_cost123bi_weibull)["scale"]

AIC(mod_cost123bi_weibull)
BIC(mod_cost123bi_weibull)

mod_cost123home_weibull<-fitdistr(cost123home,"weibull");
shapeweibull_123home<-coefficients(mod_cost123home_weibull)["shape"]
scaleweibull_123home<-coefficients(mod_cost123home_weibull)["scale"]

AIC(mod_cost123home_weibull)
BIC(mod_cost123home_weibull)

sim.t.w.cost123_md <- qweibull(sim.data.tcopula[,1], shape=shapeweibull_123md, scale = scaleweibull_123md)
sim.t.w.cost123_bi <- qweibull(sim.data.tcopula[,3],shape=shapeweibull_123bi, scale = scaleweibull_123bi)
sim.t.w.cost123_home <- qweibull(sim.data.tcopula[,2], shape=shapeweibull_123home, scale = scaleweibull_123home)

sim.t.cop.L_w=(sim.t.w.cost123_md+sim.t.w.cost123_home+sim.t.w.cost123_bi)-0.001

VaRtCopulaW<-quantile(sim.t.cop.L_w, alfa)
VaRtCopulaW

#TVaR
mean(sim.t.cop.L_w[sim.t.cop.L_w > quantile(sim.t.cop.L_w, alfa[1])])
mean(sim.t.cop.L_w[sim.t.cop.L_w > quantile(sim.t.cop.L_w, alfa[2])])
mean(sim.t.cop.L_w[sim.t.cop.L_w > quantile(sim.t.cop.L_w, alfa[3])])
cbind(mean(sim.t.cop.L_w[sim.t.cop.L_w > quantile(sim.t.cop.L_w, alfa[1])]),
mean(sim.t.cop.L_w[sim.t.cop.L_w > quantile(sim.t.cop.L_w, alfa[2])]),
mean(sim.t.cop.L_w[sim.t.cop.L_w > quantile(sim.t.cop.L_w, alfa[3])]))

#.Calcular el VaR para la pérdida L a partir de la
#cópula suponiendo marginales EXPONENCIAL con fitdist (Caso particular Weibull):

mod_cost123md_exp<-fitdistr(cost123md,"exponential");
rateexp_123md<-coefficients(mod_cost123md_exp)["rate"]

```

```

AIC(mod_cost123md_exp)
BIC(mod_cost123md_exp)

mod_cost123bi_exp<-fitdistr(cost123bi,"exponential");
rateexp_123bi<-coefficients(mod_cost123bi_exp)["rate"]

AIC(mod_cost123bi_exp)
BIC(mod_cost123bi_exp)

mod_cost123home_exp<-fitdistr(cost123home,"exponential");
rateexp_123home<-coefficients(mod_cost123home_exp)["rate"]

AIC(mod_cost123home_exp)
BIC(mod_cost123home_exp)

sim.t.E.cost123_md <- qexp(sim.data.tcopula[,1], rate = rateexp_123md)
sim.t.E.cost123_bi <- qexp(sim.data.tcopula[,3],rate = rateexp_123bi)
sim.t.E.cost123_home <- qexp(sim.data.tcopula[,2], rate = rateexp_123home)

sim.t.cop.L_E=(sim.t.E.cost123_md+sim.t.E.cost123_home+sim.t.E.cost123_bi)-0.001

VaRtCopulaE<-quantile(sim.t.cop.L_E, alfa)
VaRtCopulaE

#TVaR
mean(sim.t.cop.L_E[sim.t.cop.L_E > quantile(sim.t.cop.L_E, alfa[1])])
mean(sim.t.cop.L_E[sim.t.cop.L_E > quantile(sim.t.cop.L_E, alfa[2])])
mean(sim.t.cop.L_E[sim.t.cop.L_E > quantile(sim.t.cop.L_E, alfa[3])])
cbind(mean(sim.t.cop.L_E[sim.t.cop.L_E > quantile(sim.t.cop.L_E, alfa[1])]),
mean(sim.t.cop.L_E[sim.t.cop.L_E > quantile(sim.t.cop.L_E, alfa[2])]),
mean(sim.t.cop.L_E[sim.t.cop.L_E > quantile(sim.t.cop.L_E, alfa[3])]))

#.Calcular el VaR para la pérdida L a partir de la
#cópula suponiendo marginales GAMMA con fitdist:

mod_cost123md_gamma<-fitdistr(cost123md,"gamma");
shapegamma_123md<-coefficients(mod_cost123md_gamma)["shape"]
scalegamma_123md<-coefficients(mod_cost123md_gamma)["rate"]

AIC(mod_cost123md_gamma)
BIC(mod_cost123md_gamma)

mod_cost123bi_gamma<-fitdistr(cost123bi,"gamma");
shapegamma_123bi<-coefficients(mod_cost123bi_gamma)["shape"]
scalegamma_123bi<-coefficients(mod_cost123bi_gamma)["rate"]

AIC(mod_cost123bi_gamma)
BIC(mod_cost123bi_gamma)

mod_cost123home_gamma<-fitdistr(cost123home,"gamma");
shapegamma_123home<-coefficients(mod_cost123home_gamma)["shape"]
scalegamma_123home<-coefficients(mod_cost123home_gamma)["rate"]

AIC(mod_cost123home_gamma)
BIC(mod_cost123home_gamma)

```



```

sim.t.gam.cost123_md <- qgamma(sim.data.tcopula[,1], shape=shapegamma_123md, scale = scalegamma_123md)
sim.t.gam.cost123_home <- qgamma(sim.data.tcopula[,2], shape=shapegamma_123home, scale = scalegamma_123home)
sim.t.gam.cost123_bi <- qgamma(sim.data.tcopula[,3], shape=shapegamma_123bi, scale = scalegamma_123bi)

sim.t.cop.L_Gam=(sim.t.gam.cost123_md+sim.t.gam.cost123_home+sim.t.gam.cost123_bi)-0.001

VaRtCopulaG<-quantile(sim.t.cop.L_Gam, alfa)
VaRtCopulaG

#TVaR
mean(sim.t.cop.L_Gam[sim.t.cop.L_Gam > quantile(sim.t.cop.L_Gam, alfa[1])])
mean(sim.t.cop.L_Gam[sim.t.cop.L_Gam > quantile(sim.t.cop.L_Gam, alfa[2])])
mean(sim.t.cop.L_Gam[sim.t.cop.L_Gam > quantile(sim.t.cop.L_Gam, alfa[3])])
cbind(mean(sim.t.cop.L_Gam[sim.t.cop.L_Gam > quantile(sim.t.cop.L_Gam, alfa[1])]),
mean(sim.t.cop.L_Gam[sim.t.cop.L_Gam > quantile(sim.t.cop.L_Gam, alfa[2])]),
mean(sim.t.cop.L_Gam[sim.t.cop.L_Gam > quantile(sim.t.cop.L_Gam, alfa[3])]))

#.Calcular el VaR para la pérdida L a partir de la
#cópula suponiendo marginales Champnowne:

##Hemos de definir las funciones
pchamp = function(cost123,par)
{
a=par[1]
M=par[2]
c=par[3]
((x+c)^a-c^a)/((x+c)^a+(M+c)^a-2*c^a)
}
dchamp = function(cost123,par)
{
a=par[1]
M=par[2]
c=par[3]
(a*(x+c)^(a-1)*((M+c)^a-c^a))/((x+c)^a+(M+c)^a-2*c^a)^2
}
qchamp = function(cost123,par)
{
a=par[1]
M=par[2]
c=par[3]
((cost123*(M+c)^a+(1-2*cost123)*c^a)/(1-cost123))^(1/a)-c
}
champLike = function(cost123,par)
{
a=par[1]
M=par[2]
c=par[3]
N = length(cost123)
N*log(a)+N*log((M+c)^a-c^a)+(a-1)*sum(log(cost123+c))-
2*sum(log((cost123+c)^a+(M+c)^a-2*c^a))
}
champPar = function(cost123)
{
N = length(cost123)
par = matrix(0,nrow=3,ncol=1)
L = 0
}

```

```

avalg = 0
Lvalg = 0
cvalg = seq(0,2,0.5)
par[2] = median(cost123)
seqp = seq(1,3,1)
seqj = seq(1,21,1)
# Finder optimalt a for hver værdi af c
for (i in seq(1,length(cvalg),1))
{
par[3]=cvalg[i]*par[2]
astart=0.1
aslut=21
for (praecision in seqp)
{
for (j in seqj)
{
par[1]=astart+(j-1)*(aslut-astart)/20
L[j]=champLike(cost123,par)
}
jmax=which.max(L)
astart=max(0.0001,astart+(jmax-2)*(aslut-astart)/20)
aslut=astart+jmax*(aslut-astart)/20
}
avalg[i]=astart+(jmax-1)*(aslut-astart)/20
Lvalg[i]=L[jmax]
}
iopt=which.max(Lvalg)
par[1]=avalg[iopt]
par[3]=cvalg[iopt]*par[2]
return(par)
}

par.Ch.md123<-champPar(cost123[,1])
par.Ch.home123<-champPar(cost123[,2])
par.Ch.bi123<-champPar(cost123[,3])

AICmd123<-2*3-2*champLike(cost123[,1],par.Ch.md123)
AICbi123<-2*3-2*champLike(cost123[,3],par.Ch.bi123)
AIChome123<-2*3-2*champLike(cost123[,2],par.Ch.home123)

AICmd123
AICbi123
AIChome123

BICmd123<-3*log(n)-2*champLike(cost123[,1],par.Ch.md123)
BICbi123<-3*log(n)-2*champLike(cost123[,3],par.Ch.bi123)
BIChome123<-3*log(n)-2*champLike(cost123[,2],par.Ch.home123)

BICmd123
BICbi123
BIChome123

sim.t.Champ.cost123_md <- qchamp(sim.data.tcopula[,1], par.Ch.md123)
sim.t.Champ.cost123_home <- qchamp(sim.data.tcopula[,2], par.Ch.home123)
sim.t.Champ.cost123_bi <- qchamp(sim.data.tcopula[,3],par.Ch.bi123)

sim.t.cop.L_Champ=(sim.t.Champ.cost123_md+sim.t.Champ.cost123_home+sim.t.Champ.cost123_bi)-0.001

```

```

VaRtCopulaChamp<-quantile(sim.t.cop.L_Champ, alfa)
VaRtCopulaChamp

#TVaR
mean(sim.t.cop.L_Champ[sim.t.cop.L_Champ > quantile(sim.t.cop.L_Champ, alfa[1])])
mean(sim.t.cop.L_Champ[sim.t.cop.L_Champ > quantile(sim.t.cop.L_Champ, alfa[2])])
mean(sim.t.cop.L_Champ[sim.t.cop.L_Champ > quantile(sim.t.cop.L_Champ, alfa[3])])
cbind(mean(sim.t.cop.L_Champ[sim.t.cop.L_Champ > quantile(sim.t.cop.L_Champ, alfa[1])]),
mean(sim.t.cop.L_Champ[sim.t.cop.L_Champ > quantile(sim.t.cop.L_Champ, alfa[2])]),
mean(sim.t.cop.L_Champ[sim.t.cop.L_Champ > quantile(sim.t.cop.L_Champ, alfa[3])]))

#Cópula de referencia
sim.t.cop.Lref=(sim.t.Champ.cost123_md+sim.t.Champ.cost123_bi++sim.t.w.cost123_home)-0.001
VaRtCopulaREF<-quantile(sim.t.cop.Lref, alfa)
VaRtCopulaREF
#####CASO BIVARIANTE:BI-HOME#####

data<-read.table("Datos TFM.csv", header=TRUE, sep=";")
head(data)
costes<-cbind(data[9],data[11],data[15])
head (costes)
attach(costes)
cost<-as.matrix(costes)

cost12<-cost[(data$nclaims_md==0&data$nclaims_bi>0&data$nclaims_home>0),]
summary(cost12)

#Desviación estándar
sd(cost12[,3])
sd(cost12[,2])

#Asimetría
skewness(cost12[,3])
skewness(cost12[,2])

#Histograma
hist(cost12[,3],50, xlab = "cost_bi", main="Histograma costes daños corporales ")
hist(cost12[,2],50, xlab = "cost_home",main="Histograma costes hogar")

cost12<-((cost12/1000)+0.001)
head(cost12)
summary(cost12)
nrow(cost12)
cost12bi<-cost12[,3]
cost12home<-cost12[,2]
cost12<-cbind(cost12bi,cost12home)
head(cost12)
n<-nrow(cost12);n

###Transformación de las variables originales a variables uniformes usando la función de distribución empírica###
Ucost12<-pobs(cost12)
head(Ucost12)

###Cálculo de las correlaciones estimadas por el método clásico (correlación lineal)###

```

```

matrizpearson12<-cor(cost12)
matrizpearson12

###Estimar los coeficientes de correlación Tau y Spearman, entre los tres costes

rotau12<-Kendall(cost12)
rotau12
#roSpear<-Spearman(cost12)
#roSpear

###1.Definición parámetros iniciales###

#Parámetros iniciales-Cópula Gumbel
#Método momentos#
ParGum12<-1/(1-rotau12[1,2])
ParGum12

p12<-matrizpearson12[1,2]

#2.Definición el objeto cópula

# AJUSTE CÓPULA Y ESTIMACIÓN DE LOS PARÁMETROS DE CÓPULA #

# Cópula Gauss#
norm.cop12 <- normalCopula(p12, dim = 2, dispstr = "un")
norm.cop12
NormCopEst12<-fitCopula(norm.cop12,Ucost12, method="mpl")
NormCopEst12

# Cópula t-Student ##### Cópulas de valor extremo #
tCop12 <- tCopula(p12, dim = 2, dispstr="un", df.fixed=FALSE)
tCop12
TCopEst12<-fitCopula(tCop12, Ucost12, method="mpl", estimate.variance=TRUE)
TCopEst12

# Cópula Gumbel ##### Cópulas de valor extremo #
gumb.cop12 <- gumbelCopula(param =NA_real_, dim =2)
gumb.cop12
gumbCopEst12<-fitCopula(gumb.cop12,Ucost12, method="mpl")
gumbCopEst12

AIC(NormCopEst12)
AIC(TCopEst12)
AIC(gumbCopEst12)

BIC(NormCopEst12)
BIC(TCopEst12)
BIC(gumbCopEst12)

#mejor cópula Gaussiana

#dependencia en colas:
TailDepN<-tailIndex(norm.cop12)
TailDepN

TailDept<-tailIndex(tCop12)

```

```

TailDept

TailDepG<-tailIndex(gumb.cop12)
TailDepG

alfa<-c(0.99, 0.995,0.999)
rep=1000000

#####NORMAL#####
#####

norm.cop_est<-normalCopula(-0.4124, dim=2, dispstr = "un")
set.seed(123)
sim.data.normalcopula<-rCopula(rep,norm.cop_est)

#.Calcular el VaR para la pérdida L a partir de la
#cópula suponiendo marginales LOG-NORMAL con fitdist:

mod_cost12home_log<-fitdistr(cost12home,"lognormal");
mod_cost12home_log
meanlog_12home<-coefficients(mod_cost12home_log)["meanlog"]
sdlog_12home<-coefficients(mod_cost12home_log)["sdlog"]

AIC(mod_cost12home_log)
BIC(mod_cost12home_log)

mod_cost12bi_log<-fitdistr(cost12bi,"lognormal");
mod_cost12bi_log
meanlog_12bi<-coefficients(mod_cost12bi_log)["meanlog"]
sdlog_12bi<-coefficients(mod_cost12bi_log)["sdlog"]

AIC(mod_cost12bi_log)
BIC(mod_cost12bi_log)

sim.norm.ln.cost12_bi <- qnorm(sim.data.normalcopula[,1], mean=meanlog_12bi, sd=sdlog_12bi)
sim.norm.ln.cost12_home <- qnorm(sim.data.normalcopula[,2], mean=meanlog_12home, sd=sdlog_12home)

sim.norm.cost12_bi_log=exp(sim.norm.ln.cost12_bi)
sim.norm.cost12_home_log=exp(sim.norm.ln.cost12_home)

sim.normcop.L_log=(sim.norm.cost12_bi_log+sim.norm.cost12_home_log)-0.001

VaRnormCopulaLN<-quantile(sim.normcop.L_log, alfa)
VaRnormCopulaLN

#TVaR
mean(sim.normcop.L_log[sim.normcop.L_log > quantile(sim.normcop.L_log, alfa[1])])
mean(sim.normcop.L_log[sim.normcop.L_log > quantile(sim.normcop.L_log, alfa[2])])
mean(sim.normcop.L_log[sim.normcop.L_log > quantile(sim.normcop.L_log, alfa[3])])
cbind(mean(sim.normcop.L_log[sim.normcop.L_log > quantile(sim.normcop.L_log, alfa[1])]),
mean(sim.normcop.L_log[sim.normcop.L_log > quantile(sim.normcop.L_log, alfa[2])]),
mean(sim.normcop.L_log[sim.normcop.L_log > quantile(sim.normcop.L_log, alfa[3])]))

#.Calcular el VaR para la pérdida L a partir de la
#cópula suponiendo marginales WEIBULL con fitdist:

mod_cost12home_weibull<-fitdistr(cost12home,"weibull");

```

```

shapeweibull_12home<-coefficients(mod_cost12home_weibull)["shape"]
scaleweibull_12home<-coefficients(mod_cost12home_weibull)["scale"]

AIC(mod_cost12home_weibull)
BIC(mod_cost12home_weibull)

mod_cost12bi_weibull<-fitdistr(cost12bi,"weibull");
shapeweibull_12bi<-coefficients(mod_cost12bi_weibull)["shape"]
scaleweibull_12bi<-coefficients(mod_cost12bi_weibull)["scale"]

AIC(mod_cost12bi_weibull)
BIC(mod_cost12bi_weibull)

sim.norm.w.cost12_bi <- qweibull(sim.data.normalcopula[,1], shape=shapeweibull_12bi, scale = scaleweibull_12bi)
sim.norm.w.cost12_home <- qweibull(sim.data.normalcopula[,2],shape=shapeweibull_12home, scale = scaleweibull_12home)

sim.norm.cop.L_w=(sim.norm.w.cost12_bi+sim.norm.w.cost12_home)-0.001

VaRnormCopulaW<-quantile(sim.norm.cop.L_w, alfa)
VaRnormCopulaW

#TVaR
mean(sim.norm.cop.L_w[sim.norm.cop.L_w > quantile(sim.norm.cop.L_w, alfa[1])])
mean(sim.norm.cop.L_w[sim.norm.cop.L_w > quantile(sim.norm.cop.L_w, alfa[2])])
mean(sim.norm.cop.L_w[sim.norm.cop.L_w > quantile(sim.norm.cop.L_w, alfa[3])])
cbind(mean(sim.norm.cop.L_w[sim.norm.cop.L_w > quantile(sim.norm.cop.L_w, alfa[1])]),
mean(sim.norm.cop.L_w[sim.norm.cop.L_w > quantile(sim.norm.cop.L_w, alfa[2])]),
mean(sim.norm.cop.L_w[sim.norm.cop.L_w > quantile(sim.norm.cop.L_w, alfa[3])]))

#.Calcular el VaR para la pérdida L a partir de la
#cópula suponiendo marginales EXPONENCIAL con fitdist (Caso particular Weibull):

mod_cost12home_exp<-fitdistr(cost12home,"exponential");
rateexp_12home<-coefficients(mod_cost12home_exp)["rate"]

AIC(mod_cost12home_exp)
BIC(mod_cost12home_exp)

mod_cost12bi_exp<-fitdistr(cost12bi,"exponential");
rateexp_12bi<-coefficients(mod_cost12bi_exp)["rate"]

AIC(mod_cost12bi_exp)
BIC(mod_cost12bi_exp)

sim.gumb.E.cost12_bi <- qexp(sim.data.normalcopula[,1], rate = rateexp_12bi)
sim.gumb.E.cost12_home <- qexp(sim.data.normalcopula[,2],rate = rateexp_12home)

sim.norm.cop.L_E=(sim.gumb.E.cost12_bi+sim.gumb.E.cost12_home)-0.001

VaRnormCopula_E<-quantile(sim.norm.cop.L_E, alfa)
VaRnormCopula_E

#TVaR
mean(sim.norm.cop.L_E[sim.norm.cop.L_E > quantile(sim.norm.cop.L_E, alfa[1])])
mean(sim.norm.cop.L_E[sim.norm.cop.L_E > quantile(sim.norm.cop.L_E, alfa[2])])
mean(sim.norm.cop.L_E[sim.norm.cop.L_E > quantile(sim.norm.cop.L_E, alfa[3])])
cbind(mean(sim.norm.cop.L_E[sim.norm.cop.L_E > quantile(sim.norm.cop.L_E, alfa[1])]),

```

```

mean(sim.norm.cop.L_E[sim.norm.cop.L_E > quantile(sim.norm.cop.L_E, alfa[2])]),
mean(sim.norm.cop.L_E[sim.norm.cop.L_E > quantile(sim.norm.cop.L_E, alfa[3])]))

#.Calcular el VaR para la pérdida L a partir de la
#cópula suponiendo marginales GAMMA con fitdist:

mod_cost12home_gamma<-fitdistr(cost12home,"gamma");
shapegamma_12home<-coefficients(mod_cost12home_gamma)["shape"]
scalegamma_12home<-coefficients(mod_cost12home_gamma)["rate"]

AIC(mod_cost12home_gamma)
BIC(mod_cost12home_gamma)

mod_cost12bi_gamma<-fitdistr(cost12bi,"gamma");
shapegamma_12bi<-coefficients(mod_cost12bi_gamma)["shape"]
scalegamma_12bi<-coefficients(mod_cost12bi_gamma)["rate"]

AIC(mod_cost12bi_gamma)
BIC(mod_cost12bi_gamma)

sim.norm.gam.cost12_bi <- qgamma(sim.data.normalcopula[,1], shape=shapegamma_12bi, scale = scalegamma_12bi)
sim.norm.gam.cost12_home <- qgamma(sim.data.normalcopula[,2],shape=shapegamma_12home, scale = scalegamma_12home)

sim.norm.cop.L_Gam=(sim.norm.gam.cost12_bi+sim.norm.gam.cost12_home)-0.001

VaRnormCopulaG<-quantile(sim.norm.cop.L_Gam, alfa)
VaRnormCopulaG

#TVaR
mean(sim.norm.cop.L_Gam[sim.norm.cop.L_Gam > quantile(sim.norm.cop.L_Gam, alfa[1])])
mean(sim.norm.cop.L_Gam[sim.norm.cop.L_Gam > quantile(sim.norm.cop.L_Gam, alfa[2])])
mean(sim.norm.cop.L_Gam[sim.norm.cop.L_Gam > quantile(sim.norm.cop.L_Gam, alfa[3])])
cbind(mean(sim.norm.cop.L_Gam[sim.norm.cop.L_Gam > quantile(sim.norm.cop.L_Gam, alfa[1])]),
mean(sim.norm.cop.L_Gam[sim.norm.cop.L_Gam > quantile(sim.norm.cop.L_Gam, alfa[2])]),
mean(sim.norm.cop.L_Gam[sim.norm.cop.L_Gam > quantile(sim.norm.cop.L_Gam, alfa[3])]))

#.Calcular el VaR para la pérdida L a partir de la
#cópula suponiendo marginales Champernowne:

##Hemos de definir las funciones
pchamp = function(cost12,par)
{
a=par[1]
M=par[2]
c=par[3]
((x+c)^a-c^a)/((x+c)^a+(M+c)^a-2*c^a)
}
dchamp = function(cost12,par)
{
a=par[1]
M=par[2]
c=par[3]
(a*(x+c)^(a-1)*((M+c)^a-c^a))/((x+c)^a+(M+c)^a-2*c^a)^2
}
qchamp = function(cost12,par)
{
a=par[1]

```

```

M=par[2]
c=par[3]
((cost12*(M+c)^a+(1-2*cost12)*c^a)/(1-cost12)^(1/a)-c
}
champLike = function(cost12,par)
{
a=par[1]
M=par[2]
c=par[3]
N = length(cost12)
N*log(a)+N*log((M+c)^a-c^a)+(a-1)*sum(log(cost12+c))-
2*sum(log((cost12+c)^a+(M+c)^a-2*c^a))
}
champPar = function(cost12)
{
N = length(cost12)
par = matrix(0,nrow=3,ncol=1)
L = 0
avalg = 0
Lvalg = 0
cvalg = seq(0,2,0.5)
par[2] = median(cost12)
seqp = seq(1,3,1)
seqj = seq(1,21,1)
# Finder optimalt a for hver værdi af c
for (i in seq(1,length(cvalg),1))
{
par[3]=cvalg[i]*par[2]
astart=0.1
aslut=21
for (praecision in seqp)
{
for (j in seqj)
{
par[1]=astart+(j-1)*(aslut-astart)/20
L[j]=champLike(cost12,par)
}
jmax=which.max(L)
astart=max(0.0001,astart+(jmax-2)*(aslut-astart)/20)
aslut=astart+jmax*(aslut-astart)/20
}
avalg[i]=astart+(jmax-1)*(aslut-astart)/20
Lvalg[i]=L[jmax]
}
iopt=which.max(Lvalg)
par[1]=avalg[iopt]
par[3]=cvalg[iopt]*par[2]
return(par)
}

par.Ch.bi12<-champPar(cost12[,1])
par.Ch.home12<-champPar(cost12[,2])

AICbi12<-2*3-2*champLike(cost12[,1],par.Ch.bi12)
AIChome12<-2*3-2*champLike(cost12[,2],par.Ch.home12)

AICbi12

```



```

AIChome12

BICbi12<-3*log(n)-2*champLike(cost12[,1],par.Ch.bi12)
BIChome12<-3*log(n)-2*champLike(cost12[,2],par.Ch.home12)

BICbi12
BIChome12

sim.Norm.Champ.cost12_bi <- qchamp(sim.data.normalcopula[,1], par.Ch.bi12)
sim.Norm.Champ.cost12_home <- qchamp(sim.data.normalcopula[,2], par.Ch.home12)

sim.Norm.cop.L_Champ=(sim.Norm.Champ.cost12_bi+sim.Norm.Champ.cost12_home)-0.001

VaRNormCopulaChamp<-quantile(sim.Norm.cop.L_Champ, alfa)
VaRNormCopulaChamp

#TVaR
mean(sim.Norm.cop.L_Champ[sim.Norm.cop.L_Champ > quantile(sim.Norm.cop.L_Champ, alfa[1])])
mean(sim.Norm.cop.L_Champ[sim.Norm.cop.L_Champ > quantile(sim.Norm.cop.L_Champ, alfa[2])])
mean(sim.Norm.cop.L_Champ[sim.Norm.cop.L_Champ > quantile(sim.Norm.cop.L_Champ, alfa[3])])
cbind(mean(sim.Norm.cop.L_Champ[sim.Norm.cop.L_Champ > quantile(sim.Norm.cop.L_Champ, alfa[1])]),
mean(sim.Norm.cop.L_Champ[sim.Norm.cop.L_Champ > quantile(sim.Norm.cop.L_Champ, alfa[2])]),
mean(sim.Norm.cop.L_Champ[sim.Norm.cop.L_Champ > quantile(sim.Norm.cop.L_Champ, alfa[3])]))

sim.norm.cop.Lref=(sim.norm.w.cost12_bi+sim.Norm.Champ.cost12_home)-0.001
VaRtCopulaREF<-quantile(sim.norm.cop.Lref,alfa)
VaRtCopulaREF

##### CASO BIVARIANTE: MD-BI#####

data<-read.table("Datos TFM.csv", header=TRUE, sep=";")
head(data)
costes<-cbind(data[9],data[11],data[15])
head (costes)
attach(costes)
cost<-as.matrix(costes)

cost12<-cost[(data$nclaims_md>0&data$nclaims_bi>0&data$nclaims_home==0),]
summary(cost12)

#Desviación estándar
sd(cost12[,1])
sd(cost12[,3])

#Asimetría
skewness(cost12[,1])
skewness(cost12[,3])

#Histograma
hist(cost12[,1],50, xlab = "costes_md", main="Histograma costes daños materiales")
hist(cost12[,3],50, xlab = "costes_bi", main="Histograma costes daños corporales")

cost12<-(cost12/1000)+0.001

head(cost12)
nrow(cost12)
cost12md<-cost12[,1]

```

```

cost12bi<-cost12[,3]
cost12<-cbind(cost12md,cost12bi)
n<-nrow(cost12);n

###Transformación de las variables originales a variables uniformes usando la función de distribución empírica###
Ucost12<-pobs(cost12)
head(Ucost12)

###Cálculo de las correlaciones estimadas por el método clásico (correlación lineal)###
matrizpearson12<-cor(cost12)
matrizpearson12

###Estimar los coeficientes de correlación Tau y Spearman, entre los tres costes

rotau12<-Kendall(cost12)
rotau12

#roSpear<-Spearman(cost12)
#roSpear

###1.Definición parámetros iniciales###

#Parámetros iniciales-Cópula Gumbel
#Método momentos#
ParGum12<-1/(1-rotau12[1,2])
ParGum12

p12<-matrizpearson12[1,2]

#2.Definición el objeto cópula

# AJUSTE CÓPULA Y ESTIMACIÓN DE LOS PARÁMETROS DE CÓPULA #

# Cópula Gauss#
norm.cop12 <- normalCopula(p12, dim = 2, dispstr = "un")
norm.cop12
NormCopEst12<-fitCopula(norm.cop12,Ucost12, method="mpl")
NormCopEst12

# Cópula t-Student ##### Cópulas de valor extremo #
tCop12 <- tCopula(p12, dim = 2, dispstr="un", df.fixed=FALSE)
tCop12
TCopEst12<-fitCopula(tCop12, Ucost12, method="mpl", estimate.variance=TRUE)
TCopEst12

# Cópula Gumbel ##### Cópulas de valor extremo #
gumb.cop12 <- gumbelCopula(ParGum12, dim =2)
gumb.cop12
gumbCopEst12<-fitCopula(gumb.cop12,Ucost12, method="mpl")
gumbCopEst12

AIC(NormCopEst12)
AIC(TCopEst12)
AIC(gumbCopEst12)

BIC(NormCopEst12)
BIC(TCopEst12)

```

```

BIC(gumbCopEst12)

#mejor cópula t-Student

#dependencia en colas:
TailDepN<-tailIndex(norm.cop12)
TailDepN

TailDept<-tailIndex(tCop12)
TailDept

TailDepG<-tailIndex(gumbelCopula(ParGum12, dim =3))
TailDepG

alfa<-c(0.99, 0.995,0.999)
rep=1000000

#####T-STUDENT#####
#####

t.cop_est<-tCopula(0.1296, dim=2,dispstr = "un", df=9.1092)
set.seed(123)
sim.data.tcopula<-rCopula(rep,t.cop_est)

#.Calcular el VaR para la pérdida L a partir de la
#cópula suponiendo marginales LOG-NORMAL con fitdist:

mod_cost12md_log<-fitdistr(cost12md,"lognormal");
mod_cost12md_log
meanlog_12md<-coefficients(mod_cost12md_log)["meanlog"]
sdlog_12md<-coefficients(mod_cost12md_log)["sdlog"]

AIC(mod_cost12md_log)
BIC(mod_cost12md_log)

mod_cost12bi_log<-fitdistr(cost12bi,"lognormal");
mod_cost12bi_log
meanlog_12bi<-coefficients(mod_cost12bi_log)["meanlog"]
sdlog_12bi<-coefficients(mod_cost12bi_log)["sdlog"]

AIC(mod_cost12bi_log)
BIC(mod_cost12bi_log)

sim.t.ln.cost12_md <- qnorm(sim.data.tcopula[,1], mean=meanlog_12md, sd=sdlog_12md)
sim.t.ln.cost12_bi <- qnorm(sim.data.tcopula[,2], mean=meanlog_12bi, sd=sdlog_12bi)

sim.t.cost12_md_log=exp(sim.t.ln.cost12_md)
sim.t.cost12_bi_log=exp(sim.t.ln.cost12_bi)

sim.tcop.L_log=(sim.t.cost12_md_log+sim.t.cost12_bi_log)-0.001

VaRtCopulaLN<-quantile(sim.tcop.L_log, alfa)
VaRtCopulaLN

#TVaR
mean(sim.tcop.L_log[sim.tcop.L_log > quantile(sim.tcop.L_log, alfa[1])])
mean(sim.tcop.L_log[sim.tcop.L_log > quantile(sim.tcop.L_log, alfa[2])])

```

```

mean(sim.tcop.L_log[sim.tcop.L_log > quantile(sim.tcop.L_log, alfa[3])])
cbind(mean(sim.tcop.L_log[sim.tcop.L_log > quantile(sim.tcop.L_log, alfa[1])]),
mean(sim.tcop.L_log[sim.tcop.L_log > quantile(sim.tcop.L_log, alfa[2])]),
mean(sim.tcop.L_log[sim.tcop.L_log > quantile(sim.tcop.L_log, alfa[3])]))

#.Calcular el VaR para la pérdida L a partir de la
#cópula suponiendo marginales WEIBULL con fitdist:

mod_cost12md_weibull<-fitdistr(cost12md,"weibull");
shapeweibull_12md<-coefficients(mod_cost12md_weibull)["shape"]
scaleweibull_12md<-coefficients(mod_cost12md_weibull)["scale"]

AIC(mod_cost12md_weibull)
BIC(mod_cost12md_weibull)

mod_cost12bi_weibull<-fitdistr(cost12bi,"weibull");
shapeweibull_12bi<-coefficients(mod_cost12bi_weibull)["shape"]
scaleweibull_12bi<-coefficients(mod_cost12bi_weibull)["scale"]

AIC(mod_cost12bi_weibull)
BIC(mod_cost12bi_weibull)

sim.t.w.cost12_md <- qweibull(sim.data.tcopula[,1], shape=shapeweibull_12md, scale = scaleweibull_12md)
sim.t.w.cost12_bi <- qweibull(sim.data.tcopula[,2],shape=shapeweibull_12bi, scale = scaleweibull_12bi)

sim.t.cop.L_w=(sim.t.w.cost12_md+sim.t.w.cost12_bi)-0.001

VaRtCopulaW<-quantile(sim.t.cop.L_w, alfa)
VaRtCopulaW

#TVaR
mean(sim.t.cop.L_w[sim.t.cop.L_w > quantile(sim.t.cop.L_w, alfa[1])])
mean(sim.t.cop.L_w[sim.t.cop.L_w > quantile(sim.t.cop.L_w, alfa[2])])
mean(sim.t.cop.L_w[sim.t.cop.L_w > quantile(sim.t.cop.L_w, alfa[3])])
cbind(mean(sim.t.cop.L_w[sim.t.cop.L_w > quantile(sim.t.cop.L_w, alfa[1])]),
mean(sim.t.cop.L_w[sim.t.cop.L_w > quantile(sim.t.cop.L_w, alfa[2])]),
mean(sim.t.cop.L_w[sim.t.cop.L_w > quantile(sim.t.cop.L_w, alfa[3])]))

#.Calcular el VaR para la pérdida L a partir de la
#cópula suponiendo marginales EXPONENCIAL con fitdist (Caso particular Weibull):

mod_cost12md_exp<-fitdistr(cost12md,"exponential");
rateexp_12md<-coefficients(mod_cost12md_exp)["rate"]

AIC(mod_cost12md_exp)
BIC(mod_cost12md_exp)

mod_cost12bi_exp<-fitdistr(cost12bi,"exponential");
rateexp_12bi<-coefficients(mod_cost12bi_exp)["rate"]

AIC(mod_cost12bi_exp)
BIC(mod_cost12bi_exp)

sim.t.E.cost12_md <- qexp(sim.data.tcopula[,1], rate = rateexp_12md)
sim.t.E.cost12_bi <- qexp(sim.data.tcopula[,2],rate = rateexp_12bi)

sim.t.cop.L_E=(sim.t.E.cost12_md+sim.t.E.cost12_bi)-0.001

```

```

VaRtCopulaE<-quantile(sim.t.cop.L_E, alfa)
VaRtCopulaE

#TVaR
mean(sim.t.cop.L_E[sim.t.cop.L_E > quantile(sim.t.cop.L_E, alfa[1])])
mean(sim.t.cop.L_E[sim.t.cop.L_E > quantile(sim.t.cop.L_E, alfa[2])])
mean(sim.t.cop.L_E[sim.t.cop.L_E > quantile(sim.t.cop.L_E, alfa[3])])
cbind(mean(sim.t.cop.L_E[sim.t.cop.L_E > quantile(sim.t.cop.L_E, alfa[1])]),
mean(sim.t.cop.L_E[sim.t.cop.L_E > quantile(sim.t.cop.L_E, alfa[2])]),
mean(sim.t.cop.L_E[sim.t.cop.L_E > quantile(sim.t.cop.L_E, alfa[3])]))

#.Calcular el VaR para la pérdida L a partir de la
#cópula suponiendo marginales GAMMA con fitdistr:

mod_cost12md_gamma<-fitdistr(cost12md,"gamma");
shapegamma_12md<-coefficients(mod_cost12md_gamma)["shape"]
scalegamma_12md<-coefficients(mod_cost12md_gamma)["rate"]

AIC(mod_cost12md_gamma)
BIC(mod_cost12md_gamma)

mod_cost12bi_gamma<-fitdistr(cost12bi,"gamma");
shapegamma_12bi<-coefficients(mod_cost12bi_gamma)["shape"]
scalegamma_12bi<-coefficients(mod_cost12bi_gamma)["rate"]

AIC(mod_cost12bi_gamma)
BIC(mod_cost12bi_gamma)

sim.t.gam.cost12_md <- qgamma(sim.data.tcopula[,1], shape=shapegamma_12md, scale = scalegamma_12md)
sim.t.gam.cost12_bi <- qgamma(sim.data.tcopula[,2],shape=shapegamma_12bi, scale = scalegamma_12bi)

sim.t.cop.L_Gam=(sim.t.gam.cost12_md+sim.t.gam.cost12_bi)-0.001

VaRtCopulaG<-quantile(sim.t.cop.L_Gam, alfa)
VaRtCopulaG

#TVaR
mean(sim.t.cop.L_Gam[sim.t.cop.L_Gam > quantile(sim.t.cop.L_Gam, alfa[1])])
mean(sim.t.cop.L_Gam[sim.t.cop.L_Gam > quantile(sim.t.cop.L_Gam, alfa[2])])
mean(sim.t.cop.L_Gam[sim.t.cop.L_Gam > quantile(sim.t.cop.L_Gam, alfa[3])])
cbind(mean(sim.t.cop.L_Gam[sim.t.cop.L_Gam > quantile(sim.t.cop.L_Gam, alfa[1])]),
mean(sim.t.cop.L_Gam[sim.t.cop.L_Gam > quantile(sim.t.cop.L_Gam, alfa[2])]),
mean(sim.t.cop.L_Gam[sim.t.cop.L_Gam > quantile(sim.t.cop.L_Gam, alfa[3])]))

#.Calcular el VaR para la pérdida L a partir de la
#cópula suponiendo marginales champernowne:

##Hemos de definir las funciones
pchamp = function(cost12,par)
{
a=par[1]
M=par[2]
c=par[3]
((x+c)^a-c^a)/((x+c)^a+(M+c)^a-2*c^a)
}
dchamp = function(cost12,par)

```

```

{
a=par[1]
M=par[2]
c=par[3]
(a*(x+c)^(a-1)*((M+c)^a-c^a))/((x+c)^a+(M+c)^a-2*c^a)^2
}
qchamp = function(cost12,par)
{
a=par[1]
M=par[2]
c=par[3]
((cost12*(M+c)^a+(1-2*cost12)*c^a)/(1-cost12))^(1/a)-c
}
champLike = function(cost12,par)
{
a=par[1]
M=par[2]
c=par[3]
N = length(cost12)
N*log(a)+N*log((M+c)^a-c^a)+(a-1)*sum(log(cost12+c))-
2*sum(log((cost12+c)^a+(M+c)^a-2*c^a))
}
champPar = function(cost12)
{
N = length(cost12)
par = matrix(0,nrow=3,ncol=1)
L = 0
avalg = 0
Lvalg = 0
cvalg = seq(0,2,0.5)
par[2] = median(cost12)
seqp = seq(1,3,1)
seqj = seq(1,21,1)
# Finder optimalt a for hver værdi af c
for (i in seq(1,length(cvalg),1))
{
par[3]=cvalg[i]*par[2]
astart=0.1
aslut=21
for (praecision in seqp)
{
for (j in seqj)
{
par[1]=astart+(j-1)*(aslut-astart)/20
L[j]=champLike(cost12,par)
}
jmax=which.max(L)
astart=max(0.0001,astart+(jmax-2)*(aslut-astart)/20)
aslut=astart+jmax*(aslut-astart)/20
}
avalg[i]=astart+(jmax-1)*(aslut-astart)/20
Lvalg[i]=L[jmax]
}
iopt=which.max(Lvalg)
par[1]=avalg[iopt]
par[3]=cvalg[iopt]*par[2]
return(par)
}

```

```

}

par.Ch.md12<-champPar(cost12[,1])
par.Ch.bi12<-champPar(cost12[,2])

AICmd12<-2*3-2*champLike(cost12[,1],par.Ch.md12)
AICbi12<-2*3-2*champLike(cost12[,2],par.Ch.bi12)

AICmd12
AICbi12

BICmd12<-3*log(n)-2*champLike(cost12[,1],par.Ch.md12)
BICbi12<-3*log(n)-2*champLike(cost12[,2],par.Ch.bi12)

BICmd12
BICbi12

sim.t.Champ.cost12_md <- qchamp(sim.data.tcopula[,1], par.Ch.md12)
sim.t.Champ.cost12_bi <- qchamp(sim.data.tcopula[,2], par.Ch.bi12)

sim.t.cop.L_Champ=(sim.t.Champ.cost12_bi+sim.t.Champ.cost12_md)-0.001

VaRtCopulaChamp<-quantile(sim.t.cop.L_Champ, alfa)
VaRtCopulaChamp

#TVaR
mean(sim.t.cop.L_Champ[sim.t.cop.L_Champ > quantile(sim.t.cop.L_Champ, alfa[1])])
mean(sim.t.cop.L_Champ[sim.t.cop.L_Champ > quantile(sim.t.cop.L_Champ, alfa[2])])
mean(sim.t.cop.L_Champ[sim.t.cop.L_Champ > quantile(sim.t.cop.L_Champ, alfa[3])])
cbind(mean(sim.t.cop.L_Champ[sim.t.cop.L_Champ > quantile(sim.t.cop.L_Champ, alfa[1])]),
mean(sim.t.cop.L_Champ[sim.t.cop.L_Champ > quantile(sim.t.cop.L_Champ, alfa[2])]),
mean(sim.t.cop.L_Champ[sim.t.cop.L_Champ > quantile(sim.t.cop.L_Champ, alfa[3])]))

#VaR óptimo
sim.t.cop.Lref=(sim.t.Champ.cost12_md+sim.t.w.cost12_bi)-0.001
VaRtCopulaREF<-quantile(sim.t.cop.Lref,alfa)
VaRtCopulaREF

#####CASO BIVARIANTE:MD-HOME#####

data<-read.table("Datos TFM.csv", header=TRUE, sep=";")
head(data)
costes<-cbind(data[9],data[11],data[15])
head(costes)
attach(costes)
cost<-as.matrix(costes)

cost12<-cost[(data$nclaims_md>0&data$nclaims_bi==0&data$nclaims_home>0),]
summary(cost12)

#Desviación estándar
sd(cost12[,1])
sd(cost12[,2])

#Asimetría
skewness(cost12[,1])

```

```

skewness(cost12[,2])

#Histograma
hist(cost12[,1],50, xlab="costes_md",main="Histograma costes daños materiales")
hist(cost12[,2],50, xlab="costes_home",main="Histograma costes hogar")

cost12<-(cost12/1000)+0.001
head(cost12)
nrow(cost12)
cost12md<-cost12[,1]
cost12home<-cost12[,2]
cost12<-cbind(cost12md,cost12home)
head(cost12)
n<-nrow(cost12);n

###Transformación de las variables originales a variables uniformes usando la función de distribución empírica###
Ucost12<-pobs(cost12)
head(Ucost12)

###Cálculo de las correlaciones estimadas por el método clásico (correlación lineal)###
matrizpearson12<-cor(cost12)
matrizpearson12

###Estimar los coeficientes de correlación Tau y Spearman, entre los tres costes
rotau12<-Kendall(cost12)
rotau12

#roSpear<-Spearman(cost12)
#roSpear

###1.Definición parámetros iniciales###

#Parámetros iniciales-Cópula Gumbel
#Método momentos#
ParGum12<-1/(1-rotau12[1,2])
ParGum12

p12<-matrizpearson12[1,2]

#2.Definición el objeto cópula

# AJUSTE CÓPULA Y ESTIMACIÓN DE LOS PARÁMETROS DE CÓPULA #

# Cópula Gauss#
norm.cop12 <- normalCopula(p12, dim = 2, dispstr = "un")
norm.cop12
NormCopEst12<-fitCopula(norm.cop12,Ucost12, method="mpl")
NormCopEst12

# Cópula t-Student ##### Cópulas de valor extremo #
tCop12 <- tCopula(p12, dim = 2, dispstr="un", df.fixed=FALSE)
tCop12
TCopEst12<-fitCopula(tCop12, Ucost12, method="mpl", estimate.variance=TRUE)
TCopEst12

# Cópula Gumbel ##### Cópulas de valor extremo #
gumb.cop12 <- gumbelCopula(param =NA_real_, dim =2)

```



```

gumb.cop12
gumbCopEst12<-fitCopula(gumb.cop12,Ucost12, method="mpl")
gumbCopEst12

AIC(NormCopEst12)
AIC(TCopEst12)
AIC(gumbCopEst12)

BIC(NormCopEst12)
BIC(TCopEst12)
BIC(gumbCopEst12)

#mejor ajuste Gumbel

TailDepN<-tailIndex(norm.cop12)
TailDepN

TailDept<-tailIndex(tCop12)
TailDept

TailDepG<-tailIndex(gumb.cop12)
TailDepG

alfa<-c(0.99, 0.995,0.999)
rep=1000000

#####GUMBEL#####
#####

gumb.cop_est<-gumbelCopula(1.003, dim=2)
set.seed(123)
sim.data.gumbelcopula<-rCopula(rep,gumb.cop_est)

#.Calcular el VaR para la pérdida L a partir de la
#cópula suponiendo marginales LOG-NORMAL con fitdist:
mod_cost12md_log<-fitdistr(cost12md,"lognormal");
mod_cost12md_log
meanlog_12md<-coefficients(mod_cost12md_log)["meanlog"]
sdlog_12md<-coefficients(mod_cost12md_log)["sdlog"]

AIC(mod_cost12md_log)
BIC(mod_cost12md_log)

mod_cost12home_log<-fitdistr(cost12home,"lognormal");
mod_cost12home_log
meanlog_12home<-coefficients(mod_cost12home_log)["meanlog"]
sdlog_12home<-coefficients(mod_cost12home_log)["sdlog"]

AIC(mod_cost12home_log)
BIC(mod_cost12home_log)

sim.gumb.ln.cost12_md <- qnorm(sim.data.gumbelcopula[,1], mean=meanlog_12md, sd=sdlog_12md)
sim.gumb.ln.cost12_home <- qnorm(sim.data.gumbelcopula[,2], mean=meanlog_12home, sd=sdlog_12home)

sim.gumb.cost12_md_log=exp(sim.gumb.ln.cost12_md)
sim.gumb.cost12_home_log=exp(sim.gumb.ln.cost12_home)

```

```

sim.Gumbcop.L_log=(sim.gumb.cost12_md_log+sim.gumb.cost12_home_log)-0.001

VaRGumbCopulaLN<-quantile(sim.Gumbcop.L_log, alfa)
VaRGumbCopulaLN

#TVaR
mean(sim.Gumbcop.L_log[sim.Gumbcop.L_log > quantile(sim.Gumbcop.L_log, alfa[1])])
mean(sim.Gumbcop.L_log[sim.Gumbcop.L_log > quantile(sim.Gumbcop.L_log, alfa[2])])
mean(sim.Gumbcop.L_log[sim.Gumbcop.L_log > quantile(sim.Gumbcop.L_log, alfa[3])])
cbind(mean(sim.Gumbcop.L_log[sim.Gumbcop.L_log > quantile(sim.Gumbcop.L_log, alfa[1])]),
mean(sim.Gumbcop.L_log[sim.Gumbcop.L_log > quantile(sim.Gumbcop.L_log, alfa[2])]),
mean(sim.Gumbcop.L_log[sim.Gumbcop.L_log > quantile(sim.Gumbcop.L_log, alfa[3])]))

#. Calcular el VaR para la pérdida L a partir de la
#cópula suponiendo marginales WEIBULL con fitdist:

mod_cost12md_weibull<-fitdistr(cost12md,"weibull");
shapeweibull_12md<-coefficients(mod_cost12md_weibull) ["shape"]
scaleweibull_12md<-coefficients(mod_cost12md_weibull) ["scale"]

AIC(mod_cost12md_weibull)
BIC(mod_cost12md_weibull)

mod_cost12home_weibull<-fitdistr(cost12home,"weibull");
shapeweibull_12home<-coefficients(mod_cost12home_weibull) ["shape"]
scaleweibull_12home<-coefficients(mod_cost12home_weibull) ["scale"]

AIC(mod_cost12home_weibull)
BIC(mod_cost12home_weibull)

sim.gumb.w.cost12_md <- qweibull(sim.data.gumbelcopula[,1], shape=shapeweibull_12md, scale = scaleweibull_12md)
sim.gumb.w.cost12_home <- qweibull(sim.data.gumbelcopula[,2], shape=shapeweibull_12home, scale = scaleweibull_12home)

sim.Gumb.cop.L_w=(sim.gumb.w.cost12_md+sim.gumb.w.cost12_home)-0.001

VaRGumbCopulaW<-quantile(sim.Gumb.cop.L_w, alfa)
VaRGumbCopulaW

#TVaR
mean(sim.Gumb.cop.L_w[sim.Gumb.cop.L_w > quantile(sim.Gumb.cop.L_w, alfa[1])])
mean(sim.Gumb.cop.L_w[sim.Gumb.cop.L_w > quantile(sim.Gumb.cop.L_w, alfa[2])])
mean(sim.Gumb.cop.L_w[sim.Gumb.cop.L_w > quantile(sim.Gumb.cop.L_w, alfa[3])])
cbind(mean(sim.Gumb.cop.L_w[sim.Gumb.cop.L_w > quantile(sim.Gumb.cop.L_w, alfa[1])]),
mean(sim.Gumb.cop.L_w[sim.Gumb.cop.L_w > quantile(sim.Gumb.cop.L_w, alfa[2])]),
mean(sim.Gumb.cop.L_w[sim.Gumb.cop.L_w > quantile(sim.Gumb.cop.L_w, alfa[3])]))

#. Calcular el VaR para la pérdida L a partir de la
#cópula suponiendo marginales EXPONENCIAL con fitdist (Caso particular Weibull):

mod_cost12md_exp<-fitdistr(cost12md,"exponential");
rateexp_12md<-coefficients(mod_cost12md_exp) ["rate"]

AIC(mod_cost12md_exp)
BIC(mod_cost12md_exp)

mod_cost12home_exp<-fitdistr(cost12home,"exponential");
rateexp_12home<-coefficients(mod_cost12home_exp) ["rate"]

```

```

AIC(mod_cost12home_exp)
BIC(mod_cost12home_exp)

sim.gumb.E.cost12_md <- qexp(sim.data.gumbelcopula[,1], rate = rateexp_12md)
sim.gumb.E.cost12_home <- qexp(sim.data.gumbelcopula[,2],rate = rateexp_12home)

sim.Gumb.cop.L_E=(sim.gumb.E.cost12_md+sim.gumb.E.cost12_home)-0.001

VaRGumbCopula_E<-quantile(sim.Gumb.cop.L_E, alfa)
VaRGumbCopula_E

#TVaR
mean(sim.Gumb.cop.L_E[sim.Gumb.cop.L_E > quantile(sim.Gumb.cop.L_E, alfa[1])])
mean(sim.Gumb.cop.L_E[sim.Gumb.cop.L_E > quantile(sim.Gumb.cop.L_E, alfa[2])])
mean(sim.Gumb.cop.L_E[sim.Gumb.cop.L_E > quantile(sim.Gumb.cop.L_E, alfa[3])])
cbind(mean(sim.Gumb.cop.L_E[sim.Gumb.cop.L_E > quantile(sim.Gumb.cop.L_E, alfa[1])]),
mean(sim.Gumb.cop.L_E[sim.Gumb.cop.L_E > quantile(sim.Gumb.cop.L_E, alfa[2])]),
mean(sim.Gumb.cop.L_E[sim.Gumb.cop.L_E > quantile(sim.Gumb.cop.L_E, alfa[3])]))

#.Calcular el VaR para la pérdida L a partir de la
#cópula suponiendo marginales GAMMA con fitdist:
mod_cost12md_gamma<-fitdistr(cost12md,"gamma");
shapegamma_12md<-coefficients(mod_cost12md_gamma)["shape"]
scalegamma_12md<-coefficients(mod_cost12md_gamma)["rate"]

AIC(mod_cost12md_gamma)
BIC(mod_cost12md_gamma)

mod_cost12home_gamma<-fitdistr(cost12home,"gamma");
shapegamma_12home<-coefficients(mod_cost12home_gamma)["shape"]
scalegamma_12home<-coefficients(mod_cost12home_gamma)["rate"]

AIC(mod_cost12home_gamma)
BIC(mod_cost12home_gamma)

sim.gumb.gam.cost12_md <- qgamma(sim.data.gumbelcopula[,1], shape=shapegamma_12md, scale = scalegamma_12md)
sim.gumb.gam.cost12_home <- qgamma(sim.data.gumbelcopula[,2],shape=shapegamma_12home, scale = scalegamma_12home)

sim.Gumb.cop.L_Gam=(sim.gumb.gam.cost12_md+sim.gumb.gam.cost12_home)-0.001

VaRGumbCopulaG<-quantile(sim.Gumb.cop.L_Gam, alfa)
VaRGumbCopulaG

#TVaR
mean(sim.Gumb.cop.L_Gam[sim.Gumb.cop.L_Gam > quantile(sim.Gumb.cop.L_Gam, alfa[1])])
mean(sim.Gumb.cop.L_Gam[sim.Gumb.cop.L_Gam > quantile(sim.Gumb.cop.L_Gam, alfa[2])])
mean(sim.Gumb.cop.L_Gam[sim.Gumb.cop.L_Gam > quantile(sim.Gumb.cop.L_Gam, alfa[3])])
cbind(mean(sim.Gumb.cop.L_Gam[sim.Gumb.cop.L_Gam > quantile(sim.Gumb.cop.L_Gam, alfa[1])]),
mean(sim.Gumb.cop.L_Gam[sim.Gumb.cop.L_Gam > quantile(sim.Gumb.cop.L_Gam, alfa[2])]),
mean(sim.Gumb.cop.L_Gam[sim.Gumb.cop.L_Gam > quantile(sim.Gumb.cop.L_Gam, alfa[3])]))

#.Calcular el VaR para la pérdida L a partir de la
#cópula suponiendo marginales champernowne:
#Champernowne
##Hemos de definir las funciones
pchamp = function(cost12,par)

```

```

{
a=par[1]
M=par[2]
c=par[3]
((x+c)^a-c^a)/((x+c)^a+(M+c)^a-2*c^a)
}
dchamp = function(cost12,par)
{
a=par[1]
M=par[2]
c=par[3]
(a*(x+c)^(a-1)*((M+c)^a-c^a))/((x+c)^a+(M+c)^a-2*c^a)^2
}
qchamp = function(cost12,par)
{
a=par[1]
M=par[2]
c=par[3]
((cost12*(M+c)^a+(1-2*cost12)*c^a)/(1-cost12))^(1/a)-c
}
champLike = function(cost12,par)
{
a=par[1]
M=par[2]
c=par[3]
N = length(cost12)
N*log(a)+N*log((M+c)^a-c^a)+(a-1)*sum(log(cost12+c))-
2*sum(log((cost12+c)^a+(M+c)^a-2*c^a))
}
champPar = function(cost12)
{
N = length(cost12)
par = matrix(0,nrow=3,ncol=1)
L = 0
avalg = 0
lvalg = 0
cvalg = seq(0,2,0.5)
par[2] = median(cost12)
seqp = seq(1,3,1)
seqj = seq(1,21,1)
# Finder optimalt a for hver værdi af c
for (i in seq(1,length(cvalg),1))
{
par[3]=cvalg[i]*par[2]
astart=0.1
aslut=21
for (praecision in seqp)
{
for (j in seqj)
{
par[1]=astart+(j-1)*(aslut-astart)/20
L[j]=champLike(cost12,par)
}
jmax=which.max(L)
astart=max(0.0001,astart+(jmax-2)*(aslut-astart)/20)
aslut=astart+jmax*(aslut-astart)/20
}
}
}

```

```

    avalg[i]=astart+(jmax-1)*(aslut-astart)/20
    Lvalg[i]=L[jmax]
  }
  iopt=which.max(Lvalg)
  par[1]=avalg[iopt]
  par[3]=cvalg[iopt]*par[2]
  return(par)
}

par.Ch.md12<-champPar(cost12[,1])
par.Ch.home12<-champPar(cost12[,2])

AICmd12<-2*3-2*champLike(cost12[,1],par.Ch.md12)
AIChome12<-2*3-2*champLike(cost12[,2],par.Ch.home12)

AICmd12
AIChome12

BICmd12<-3*log(n)-2*champLike(cost12[,1],par.Ch.md12)
BIChome12<-3*log(n)-2*champLike(cost12[,2],par.Ch.home12)

BICmd12
BIChome12

sim.Gum.Champ.cost12_md <- qchamp(sim.data.gumbelcopula[,1], par.Ch.md12)
sim.Gum.Champ.cost12_home <- qchamp(sim.data.gumbelcopula[,2], par.Ch.home12)

sim.Gum.cop.L_Champ=(sim.Gum.Champ.cost12_md+sim.Gum.Champ.cost12_home)-0.001

VaRGumCopulaChamp<-quantile(sim.Gum.cop.L_Champ, alfa)
VaRGumCopulaChamp

#TVaR
mean(sim.Gum.cop.L_Champ[sim.Gum.cop.L_Champ > quantile(sim.Gum.cop.L_Champ, alfa[1])])
mean(sim.Gum.cop.L_Champ[sim.Gum.cop.L_Champ > quantile(sim.Gum.cop.L_Champ, alfa[2])])
mean(sim.Gum.cop.L_Champ[sim.Gum.cop.L_Champ > quantile(sim.Gum.cop.L_Champ, alfa[3])])
cbind(mean(sim.Gum.cop.L_Champ[sim.Gum.cop.L_Champ > quantile(sim.Gum.cop.L_Champ, alfa[1])]),
mean(sim.Gum.cop.L_Champ[sim.Gum.cop.L_Champ > quantile(sim.Gum.cop.L_Champ, alfa[2])]),
mean(sim.Gum.cop.L_Champ[sim.Gum.cop.L_Champ > quantile(sim.Gum.cop.L_Champ, alfa[3])]))

#VaR óptimo:
sim.Gumb.cop.Lref=(sim.Gum.Champ.cost12_md+sim.gumb.w.cost12_home)-0.001
VaRtCopulaREF<-quantile(sim.Gumb.cop.Lref,alfa)
VaRtCopulaREF
#####CASO UNIVARIANTE:MD #####

data<-read.table("Datos TFM.csv", header=TRUE, sep=";")
head(data)
costes<-cbind(data[9],data[11],data[15])
head(costes)
attach(costes)
cost<-as.matrix(costes)
nrow(costes)

cost1md<-cost[(data$nclaims_md>0&data$nclaims_bi==0&data$nclaims_home==0),]
head(cost1md)
nrow(cost1md)

```

```

n<-nrow(cost1md)
cost1md<-cost1md[,1]
summary(cost1md)
sd(cost1md)
skewness(cost1md)

hist(cost1md,50, xlab="coste_md",main="Histograma costes daños materiales")

alfa<-c(0.99, 0.995,0.999)
rep=1000000

set.seed(123)
sim.data<-runif(rep)

cost1md<-(cost1md/1000)+0.001

#.Calcular el VaR para la pérdida L suponiendo marginales LOG-NORMAL con fitdist:
##Cálculo de los parámetros:

mod_cost1md_log<-fitdistr(cost1md,"lognormal");
mod_cost1md_log
meanlog_1md<-coefficients(mod_cost1md_log)["meanlog"]
sdlog_1md<-coefficients(mod_cost1md_log)["sdlog"]

AIC(mod_cost1md_log)
BIC(mod_cost1md_log)

sim.data.cost1md.log<- rlnorm(rep,meanlog_1md,sdlog_1md)-0.001

VaRLN.md<-quantile(sim.data.cost1md.log, alfa)
VaRLN.md

#TVaR
mean(sim.data.cost1md.log[sim.data.cost1md.log > quantile(sim.data.cost1md.log, alfa[1])])
mean(sim.data.cost1md.log[sim.data.cost1md.log > quantile(sim.data.cost1md.log, alfa[2])])
mean(sim.data.cost1md.log[sim.data.cost1md.log > quantile(sim.data.cost1md.log, alfa[3])])
cbind(mean(sim.data.cost1md.log[sim.data.cost1md.log > quantile(sim.data.cost1md.log, alfa[1])]),
mean(sim.data.cost1md.log[sim.data.cost1md.log > quantile(sim.data.cost1md.log, alfa[2])]),
mean(sim.data.cost1md.log[sim.data.cost1md.log > quantile(sim.data.cost1md.log, alfa[3])]))

#.Calcular el VaR para la pérdida L suponiendo marginales WEIBULL con fitdist:
##Cálculo de los parámetros:

mod_cost1md_w<-fitdistr(cost1md,"weibull");
mod_cost1md_w
shapew_1md<-coefficients(mod_cost1md_w)["shape"]
scalew_1md<-coefficients(mod_cost1md_w)["scale"]

AIC(mod_cost1md_w)
BIC(mod_cost1md_w)

sim.data.cost1md.w<- rweibull(rep,shapew_1md,scalew_1md)-0.001

VaRW.md<-quantile(sim.data.cost1md.w, alfa)
VaRW.md

#TVaR

```

```

mean(sim.data.cost1md.w[sim.data.cost1md.w > quantile(sim.data.cost1md.w, alfa[1])])
mean(sim.data.cost1md.w[sim.data.cost1md.w > quantile(sim.data.cost1md.w, alfa[2])])
mean(sim.data.cost1md.w[sim.data.cost1md.w > quantile(sim.data.cost1md.w, alfa[3])])
cbind(mean(sim.data.cost1md.w[sim.data.cost1md.w > quantile(sim.data.cost1md.w, alfa[1])]),
mean(sim.data.cost1md.w[sim.data.cost1md.w > quantile(sim.data.cost1md.w, alfa[2])]),
mean(sim.data.cost1md.w[sim.data.cost1md.w > quantile(sim.data.cost1md.w, alfa[3])]))

#.Calcular el VaR para la pérdida L suponiendo marginales EXPONENCIAL con fitdist:
##Cálculo de los parámetros:

mod_cost1md_exp<-fitdistr(cost1md,"exponential");
rateexp_1md<-coefficients(mod_cost1md_exp)["rate"]

AIC(mod_cost1md_exp)
BIC(mod_cost1md_exp)

sim.data.cost1md.exp<- rexp(rep,rate=rateexp_1md)-0.001
VaRExp.md<-quantile(sim.data.cost1md.exp, alfa)
VaRExp.md

#TVaR
mean(sim.data.cost1md.exp[sim.data.cost1md.exp > quantile(sim.data.cost1md.exp, alfa[1])])
mean(sim.data.cost1md.exp[sim.data.cost1md.exp > quantile(sim.data.cost1md.exp, alfa[2])])
mean(sim.data.cost1md.exp[sim.data.cost1md.exp > quantile(sim.data.cost1md.exp, alfa[3])])
cbind(mean(sim.data.cost1md.exp[sim.data.cost1md.exp > quantile(sim.data.cost1md.exp, alfa[1])]),
mean(sim.data.cost1md.exp[sim.data.cost1md.exp > quantile(sim.data.cost1md.exp, alfa[2])]),
mean(sim.data.cost1md.exp[sim.data.cost1md.exp > quantile(sim.data.cost1md.exp, alfa[3])]))

#.Calcular el VaR para la pérdida L suponiendo marginales GAMMA con fitdist:
##Cálculo de los parámetros:

mod_cost1md_gamma<-fitdistr(cost1md,"gamma");
shapegamma_1md<-coefficients(mod_cost1md_gamma)["shape"]
scalegamma_1md<-coefficients(mod_cost1md_gamma)["rate"]

AIC(mod_cost1md_gamma)
BIC(mod_cost1md_gamma)

sim.data.cost1md.gamma<- rgamma(rep,shapegamma_1md,scalegamma_1md)-0.001
VaRGam.md<-quantile(sim.data.cost1md.gamma, alfa)
VaRGam.md

#TvaR
mean(sim.data.cost1md.gamma[sim.data.cost1md.gamma > quantile(sim.data.cost1md.gamma, alfa[1])])
mean(sim.data.cost1md.gamma[sim.data.cost1md.gamma > quantile(sim.data.cost1md.gamma, alfa[2])])
mean(sim.data.cost1md.gamma[sim.data.cost1md.gamma > quantile(sim.data.cost1md.gamma, alfa[3])])
cbind(mean(sim.data.cost1md.gamma[sim.data.cost1md.gamma > quantile(sim.data.cost1md.gamma, alfa[1])]),
mean(sim.data.cost1md.gamma[sim.data.cost1md.gamma > quantile(sim.data.cost1md.gamma, alfa[2])]),
mean(sim.data.cost1md.gamma[sim.data.cost1md.gamma > quantile(sim.data.cost1md.gamma, alfa[3])]))

#.Calcular el VaR para la pérdida L suponiendo marginales CHAMPERNOWNE con fitdist:
##Cálculo de los parámetros:
pchamp = function(cost1md,par)
{
a=par[1]
M=par[2]
c=par[3]

```

```

((x+c)^a-c^a)/((x+c)^a+(M+c)^a-2*c^a)
}
dchamp = function(cost1md,par)
{
a=par[1]
M=par[2]
c=par[3]
(a*(x+c)^(a-1)*((M+c)^a-c^a))/((x+c)^a+(M+c)^a-2*c^a)^2
}
qchamp = function(cost1md,par)
{
a=par[1]
M=par[2]
c=par[3]
((cost1md*(M+c)^a+(1-2*cost1md)*c^a)/(1-cost1md))^(1/a)-c
}
champLike = function(cost1md,par)
{
a=par[1]
M=par[2]
c=par[3]
N = length(cost1md)
N*log(a)+N*log((M+c)^a-c^a)+(a-1)*sum(log(cost1md+c))-
2*sum(log((cost1md+c)^a+(M+c)^a-2*c^a))
}
champPar = function(cost1md)
{
N = length(cost1md)
par = matrix(0,nrow=3,ncol=1)
L = 0
avalg = 0
Lvalg = 0
cvalg = seq(0,2,0.5)
par[2] = median(cost1md)
seqp = seq(1,3,1)
seqj = seq(1,21,1)
# Finder optimalt a for hver værdi af c
for (i in seq(1,length(cvalg),1))
{
par[3]=cvalg[i]*par[2]
astart=0.1
aslut=21
for (praecision in seqp)
{
for (j in seqj)
{
par[1]=astart+(j-1)*(aslut-astart)/20
L[j]=champLike(cost1md,par)
}
jmax=which.max(L)
astart=max(0.0001,astart+(jmax-2)*(aslut-astart)/20)
aslut=astart+jmax*(aslut-astart)/20
}
avalg[i]=astart+(jmax-1)*(aslut-astart)/20
Lvalg[i]=L[jmax]
}
iopt=which.max(Lvalg)
}

```



```

par[1]=avalg[iopt]
par[3]=cvalg[iopt]*par[2]
return(par)
}

par.Ch.md<-champPar(cost1md)

AICmd12<-2*3-2*champLike(cost1md,par.Ch.md)
AICmd12

BICmd12<-3*log(n)-2*champLike(cost1md,par.Ch.md)
BICmd12

sim.Champ.cost1_md <- qchamp(sim.data, par.Ch.md)
sim.L_Champ=sim.Champ.cost1_md-0.001

VaRChamp<-quantile(sim.L_Champ, alfa)
VaRChamp

#TVaR
mean(sim.L_Champ[sim.L_Champ > quantile(sim.L_Champ, alfa[1])])
mean(sim.L_Champ[sim.L_Champ > quantile(sim.L_Champ, alfa[2])])
mean(sim.L_Champ[sim.L_Champ > quantile(sim.L_Champ, alfa[3])])
cbind(mean(sim.L_Champ[sim.L_Champ > quantile(sim.L_Champ, alfa[1])]),
mean(sim.L_Champ[sim.L_Champ > quantile(sim.L_Champ, alfa[2])]),
mean(sim.L_Champ[sim.L_Champ > quantile(sim.L_Champ, alfa[3])]))

##### CASO UNIVARIANTE:BI#####

data<-read.table("Datos TFM.csv", header=TRUE, sep=";")
head(data)
costes<-cbind(data[9],data[11],data[15])
head (costes)
attach(costes)
cost<-as.matrix(costes)
nrow(costes)

cost1bi<-cost[(data$nclaims_md==0&data$nclaims_bi>0&data$nclaims_home==0),]
head(cost1bi)
nrow(cost1bi)
n<-nrow(cost1bi);n
summary(cost1bi)
cost1bi<-cost1bi[,3]
sd(cost1bi)
skewness(cost1bi)
hist(cost1bi,50,xlab="costes_bi" ,main="Histograma costes daños corporales")

alfa<-c(0.99, 0.995,0.999)
rep=1000000

set.seed(123)
sim.data<-runif(rep)

cost1bi<-(cost1bi/1000)+0.001

#. Calcular el VaR para la pérdida L suponiendo marginales LOG-NORMAL con fitdist:
##Cálculo de los parámetros:

```

```

mod_cost1bi_log<-fitdistr(cost1bi,"lognormal");
mod_cost1bi_log
meanlog_1bi<-coefficients(mod_cost1bi_log)["meanlog"]
sdlog_1bi<-coefficients(mod_cost1bi_log)["sdlog"]

AIC(mod_cost1bi_log)
BIC(mod_cost1bi_log)

sim.data.cost1bi.log<- rlnorm(rep,meanlog_1bi,sdlog_1bi)-0.001

VaRLN.bi<-quantile(sim.data.cost1bi.log, alfa)
VaRLN.bi

#TVaR
mean(sim.data.cost1bi.log[sim.data.cost1bi.log > quantile(sim.data.cost1bi.log, alfa[1])])
mean(sim.data.cost1bi.log[sim.data.cost1bi.log > quantile(sim.data.cost1bi.log, alfa[2])])
mean(sim.data.cost1bi.log[sim.data.cost1bi.log > quantile(sim.data.cost1bi.log, alfa[3])])
cbind(mean(sim.data.cost1bi.log[sim.data.cost1bi.log > quantile(sim.data.cost1bi.log, alfa[1])]),
mean(sim.data.cost1bi.log[sim.data.cost1bi.log > quantile(sim.data.cost1bi.log, alfa[2])]),
mean(sim.data.cost1bi.log[sim.data.cost1bi.log > quantile(sim.data.cost1bi.log, alfa[3])]))

#. Calcular el VaR para la pérdida L suponiendo marginales WEIBULL con fitdist:
##Cálculo de los parámetros:

mod_cost1bi_w<-fitdistr(cost1bi,"weibull");
mod_cost1bi_w
shapew_1bi<-coefficients(mod_cost1bi_w)["shape"]
scalew_1bi<-coefficients(mod_cost1bi_w)["scale"]

AIC(mod_cost1bi_w)
BIC(mod_cost1bi_w)

sim.data.cost1bi.w<- rweibull(rep,shapew_1bi,scalew_1bi)-0.001

VaRW.bi<-quantile(sim.data.cost1bi.w, alfa)
VaRW.bi
#TVaR
mean(sim.data.cost1bi.w[sim.data.cost1bi.w > quantile(sim.data.cost1bi.w, alfa[1])])
mean(sim.data.cost1bi.w[sim.data.cost1bi.w > quantile(sim.data.cost1bi.w, alfa[2])])
mean(sim.data.cost1bi.w[sim.data.cost1bi.w > quantile(sim.data.cost1bi.w, alfa[3])])
cbind(mean(sim.data.cost1bi.w[sim.data.cost1bi.w > quantile(sim.data.cost1bi.w, alfa[1])]),
mean(sim.data.cost1bi.w[sim.data.cost1bi.w > quantile(sim.data.cost1bi.w, alfa[2])]),
mean(sim.data.cost1bi.w[sim.data.cost1bi.w > quantile(sim.data.cost1bi.w, alfa[3])]))

#. Calcular el VaR para la pérdida L suponiendo marginales EXPONENCIAL con fitdist:
##Cálculo de los parámetros:

mod_cost1bi_exp<-fitdistr(cost1bi,"exponential");
rateexp_1bi<-coefficients(mod_cost1bi_exp)["rate"]

AIC(mod_cost1bi_exp)
BIC(mod_cost1bi_exp)

sim.data.cost1bi.exp<- rexp(rep,rate=rateexp_1bi)-0.001
VaRExp.bi<-quantile(sim.data.cost1bi.exp, alfa)
VaRExp.bi

```

```

#TVaR
mean(sim.data.cost1bi.exp[sim.data.cost1bi.exp > quantile(sim.data.cost1bi.exp, alfa[1])])
mean(sim.data.cost1bi.exp[sim.data.cost1bi.exp > quantile(sim.data.cost1bi.exp, alfa[2])])
mean(sim.data.cost1bi.exp[sim.data.cost1bi.exp > quantile(sim.data.cost1bi.exp, alfa[3])])
cbind(mean(sim.data.cost1bi.exp[sim.data.cost1bi.exp > quantile(sim.data.cost1bi.exp, alfa[1])]),
mean(sim.data.cost1bi.exp[sim.data.cost1bi.exp > quantile(sim.data.cost1bi.exp, alfa[2])]),
mean(sim.data.cost1bi.exp[sim.data.cost1bi.exp > quantile(sim.data.cost1bi.exp, alfa[3])]))

#.Calcular el VaR para la pérdida L suponiendo marginales GAMMA con fitdist:
##Cálculo de los parámetros:

mod_cost1bi_gamma<-fitdistr(cost1bi,"gamma");
shapegamma_1bi<-coefficients(mod_cost1bi_gamma)["shape"]
scalegamma_1bi<-coefficients(mod_cost1bi_gamma)["rate"]

AIC(mod_cost1bi_gamma)
BIC(mod_cost1bi_gamma)

sim.data.cost1bi.gamma<- rgamma(rep,shapegamma_1bi,scalegamma_1bi)-0.001
VaRGam.bi<-quantile(sim.data.cost1bi.gamma, alfa)
VaRGam.bi

#TvaR
mean(sim.data.cost1bi.gamma[sim.data.cost1bi.gamma > quantile(sim.data.cost1bi.gamma, alfa[1])])
mean(sim.data.cost1bi.gamma[sim.data.cost1bi.gamma > quantile(sim.data.cost1bi.gamma, alfa[2])])
mean(sim.data.cost1bi.gamma[sim.data.cost1bi.gamma > quantile(sim.data.cost1bi.gamma, alfa[3])])
cbind(mean(sim.data.cost1bi.gamma[sim.data.cost1bi.gamma > quantile(sim.data.cost1bi.gamma, alfa[1])]),
mean(sim.data.cost1bi.gamma[sim.data.cost1bi.gamma > quantile(sim.data.cost1bi.gamma, alfa[2])]),
mean(sim.data.cost1bi.gamma[sim.data.cost1bi.gamma > quantile(sim.data.cost1bi.gamma, alfa[3])]))

#.Calcular el VaR para la pérdida L suponiendo marginales CHAMPERNOWNE con fitdist:
##Cálculo de los parámetros:
pchamp = function(cost1bi,par)
{
a=par[1]
M=par[2]
c=par[3]
((x+c)^a-c^a)/((x+c)^a+(M+c)^a-2*c^a)
}
dchamp = function(cost1bi,par)
{
a=par[1]
M=par[2]
c=par[3]
(a*(x+c)^(a-1)*((M+c)^a-c^a))/((x+c)^a+(M+c)^a-2*c^a)^2
}
qchamp = function(cost1bi,par)
{
a=par[1]
M=par[2]
c=par[3]
((cost1bi*(M+c)^a+(1-2*cost1bi)*c^a)/(1-cost1bi))^(1/a)-c
}
champLike = function(cost1bi,par)
{
a=par[1]

```

```

M=par[2]
c=par[3]
N = length(cost1bi)
N*log(a)+N*log((M+c)^a-c^a)+(a-1)*sum(log(cost1bi+c))-
2*sum(log((cost1bi+c)^a+(M+c)^a-2*c^a))
}
champPar = function(cost1bi)
{
N = length(cost1bi)
par = matrix(0,nrow=3,ncol=1)
L = 0
avalg = 0
Lvalg = 0
cvalg = seq(0,2,0.5)
par[2] = median(cost1bi)
seqp = seq(1,3,1)
seqj = seq(1,21,1)
# Finder optimalt a for hver værdi af c
for (i in seq(1,length(cvalg),1))
{
par[3]=cvalg[i]*par[2]
astart=0.1
aslut=21
for (praecision in seqp)
{
for (j in seqj)
{
par[1]=astart+(j-1)*(aslut-astart)/20
L[j]=champLike(cost1bi,par)
}
jmax=which.max(L)
astart=max(0.0001,astart+(jmax-2)*(aslut-astart)/20)
aslut=astart+jmax*(aslut-astart)/20
}
avalg[i]=astart+(jmax-1)*(aslut-astart)/20
Lvalg[i]=L[jmax]
}
iopt=which.max(Lvalg)
par[1]=avalg[iopt]
par[3]=cvalg[iopt]*par[2]
return(par)
}
par.Ch.bi<-champPar(cost1bi)
AICmd12<-2*3-2*champLike(cost1bi,par.Ch.bi)
AICmd12

BICmd12<-3*log(n)-2*champLike(cost1bi,par.Ch.bi)
BICmd12

sim.Champ.cost1_bi <- qchamp(sim.data, par.Ch.bi)
sim.L_Champ=sim.Champ.cost1_bi-0.001

VaRChamp<-quantile(sim.L_Champ, alfa)
VaRChamp

#TVaR
mean(sim.L_Champ[sim.L_Champ > quantile(sim.L_Champ, alfa[1])])

```

```

mean(sim.L_Champ[sim.L_Champ > quantile(sim.L_Champ, alfa[2])])
mean(sim.L_Champ[sim.L_Champ > quantile(sim.L_Champ, alfa[3])])
cbind(mean(sim.L_Champ[sim.L_Champ > quantile(sim.L_Champ, alfa[1])]),
mean(sim.L_Champ[sim.L_Champ > quantile(sim.L_Champ, alfa[2])]),
mean(sim.L_Champ[sim.L_Champ > quantile(sim.L_Champ, alfa[3])]))

##### CASO UNIVARIANTE:HOME#####

data<-read.table("Datos TFM.csv", header=TRUE, sep=";")
head(data)
costes<-cbind(data[9],data[11],data[15])
head (costes)
attach(costes)
cost<-as.matrix(costes)
nrow(costes)

cost1home<-cost[(data$nclaims_md==0&data$nclaims_bi==0&data$nclaims_home>0),]
head(cost1home)
nrow(cost1home)
n<-nrow(cost1home);n
cost1home<-cost1home[,2]
summary(cost1home)
sd(cost1home)
skewness(cost1home)
hist(cost1home,50, xlab = "costes_hogar", main="Histogramas costes hogar")

alfa<-c(0.99, 0.995,0.999)
rep=1000000

set.seed(123)
sim.data<-runif(rep)

cost1home<-(cost1home/1000)+0.001

#.Calcular el VaR para la pérdida L suponiendo marginales LOG-NORMAL con fitdist:
##Cálculo de los parámetros:

mod_cost1home_log<-fitdistr(cost1home,"lognormal");
mod_cost1home_log
meanlog_1home<-coefficients(mod_cost1home_log)["meanlog"]
sdlog_1home<-coefficients(mod_cost1home_log)["sdlog"]

AIC(mod_cost1home_log)
BIC(mod_cost1home_log)

sim.data.cost1home.log<- rlnorm(rep,meanlog_1home,sdlog_1home)-0.001

VaRLN.home<-quantile(sim.data.cost1home.log, alfa)
VaRLN.home

#TVaR
mean(sim.data.cost1home.log[sim.data.cost1home.log > quantile(sim.data.cost1home.log, alfa[1])])
mean(sim.data.cost1home.log[sim.data.cost1home.log > quantile(sim.data.cost1home.log, alfa[2])])
mean(sim.data.cost1home.log[sim.data.cost1home.log > quantile(sim.data.cost1home.log, alfa[3])])
cbind(mean(sim.data.cost1home.log[sim.data.cost1home.log > quantile(sim.data.cost1home.log, alfa[1])]),

```

```

mean(sim.data.cost1home.log[sim.data.cost1home.log > quantile(sim.data.cost1home.log, alfa[2])]),
mean(sim.data.cost1home.log[sim.data.cost1home.log > quantile(sim.data.cost1home.log, alfa[3])]))

#.Calcular el VaR para la pérdida L suponiendo marginales WEIBULL con fitdist:
##Cálculo de los parámetros:

mod_cost1home_w<-fitdistr(cost1home,"weibull");
mod_cost1home_w
shapew_1home<-coefficients(mod_cost1home_w)["shape"]
scalew_1home<-coefficients(mod_cost1home_w)["scale"]

AIC(mod_cost1home_w)
BIC(mod_cost1home_w)

sim.data.cost1home.w<- rweibull(rep,shapew_1home,scalew_1home)-0.001

VaRW.home<-quantile(sim.data.cost1home.w, alfa)
VaRW.home

#TVaR
mean(sim.data.cost1home.w[sim.data.cost1home.w > quantile(sim.data.cost1home.w, alfa[1])])
mean(sim.data.cost1home.w[sim.data.cost1home.w > quantile(sim.data.cost1home.w, alfa[2])])
mean(sim.data.cost1home.w[sim.data.cost1home.w > quantile(sim.data.cost1home.w, alfa[3])])
cbind(mean(sim.data.cost1home.w[sim.data.cost1home.w > quantile(sim.data.cost1home.w, alfa[1])]),
mean(sim.data.cost1home.w[sim.data.cost1home.w > quantile(sim.data.cost1home.w, alfa[2])]),
mean(sim.data.cost1home.w[sim.data.cost1home.w > quantile(sim.data.cost1home.w, alfa[3])]))

#.Calcular el VaR para la pérdida L suponiendo marginales EXPONENCIAL con fitdist:
##Cálculo de los parámetros:

mod_cost1home_exp<-fitdistr(cost1home,"exponential");
rateexp_1home<-coefficients(mod_cost1home_exp)["rate"]

AIC(mod_cost1home_exp)
BIC(mod_cost1home_exp)

sim.data.cost1home.exp<- rexp(rep,rate=rateexp_1home)-0.001
VaRExp.home<-quantile(sim.data.cost1home.exp, alfa)
VaRExp.home

#TVaR
mean(sim.data.cost1home.exp[sim.data.cost1home.exp > quantile(sim.data.cost1home.exp, alfa[1])])
mean(sim.data.cost1home.exp[sim.data.cost1home.exp > quantile(sim.data.cost1home.exp, alfa[2])])
mean(sim.data.cost1home.exp[sim.data.cost1home.exp > quantile(sim.data.cost1home.exp, alfa[3])])
cbind(mean(sim.data.cost1home.exp[sim.data.cost1home.exp > quantile(sim.data.cost1home.exp, alfa[1])]),
mean(sim.data.cost1home.exp[sim.data.cost1home.exp > quantile(sim.data.cost1home.exp, alfa[2])]),
mean(sim.data.cost1home.exp[sim.data.cost1home.exp > quantile(sim.data.cost1home.exp, alfa[3])]))

#.Calcular el VaR para la pérdida L suponiendo marginales GAMMA con fitdist:
##Cálculo de los parámetros:

mod_cost1home_gamma<-fitdistr(cost1home,"gamma");
shapegamma_1home<-coefficients(mod_cost1home_gamma)["shape"]
scalegamma_1home<-coefficients(mod_cost1home_gamma)["rate"]

```

```

AIC(mod_cost1home_gamma)
BIC(mod_cost1home_gamma)

sim.data.cost1home.gamma<- rgamma(rep,shapegamma_1home,scalegamma_1home)-0.001
VaRGam.home<-quantile(sim.data.cost1home.gamma, alfa)
VaRGam.home

#TvaR
mean(sim.data.cost1home.gamma[sim.data.cost1home.gamma > quantile(sim.data.cost1home.gamma, alfa[1])])
mean(sim.data.cost1home.gamma[sim.data.cost1home.gamma > quantile(sim.data.cost1home.gamma, alfa[2])])
mean(sim.data.cost1home.gamma[sim.data.cost1home.gamma > quantile(sim.data.cost1home.gamma, alfa[3])])
cbind(mean(sim.data.cost1home.gamma[sim.data.cost1home.gamma > quantile(sim.data.cost1home.gamma, alfa[1])])
,mean(sim.data.cost1home.gamma[sim.data.cost1home.gamma > quantile(sim.data.cost1home.gamma, alfa[2])]),
mean(sim.data.cost1home.gamma[sim.data.cost1home.gamma > quantile(sim.data.cost1home.gamma, alfa[3])]))

#. Calcular el VaR para la pérdida L suponiendo marginales CHAMPERNOWNE con fitdist:
##Cálculo de los parámetros:

pchamp = function(cost1home,par)
{
a=par[1]
M=par[2]
c=par[3]
((x+c)^a-c^a)/((x+c)^a+(M+c)^a-2*c^a)
}
dchamp = function(cost1home,par)
{
a=par[1]
M=par[2]
c=par[3]
(a*(x+c)^(a-1)*((M+c)^a-c^a))/((x+c)^a+(M+c)^a-2*c^a)^2
}
qchamp = function(cost1home,par)
{
a=par[1]
M=par[2]
c=par[3]
((cost1home*(M+c)^a+(1-2*cost1home)*c^a)/(1-cost1home))^(1/a)-c
}
champLike = function(cost1home,par)
{
a=par[1]
M=par[2]
c=par[3]
N = length(cost1home)
N*log(a)+N*log((M+c)^a-c^a)+(a-1)*sum(log(cost1home+c))-
2*sum(log((cost1home+c)^a+(M+c)^a-2*c^a))
}
champPar = function(cost1home)
{
N = length(cost1home)
par = matrix(0,nrow=3,ncol=1)
L = 0
avalg = 0
Lvalg = 0
cvalg = seq(0,2,0.5)
par[2] = median(cost1home)
}

```

```

seqp = seq(1,3,1)
seqj = seq(1,21,1)
# Finder optimalt a for hver værdi af c
for (i in seq(1,length(cvalg),1))
{
par[3]=cvalg[i]*par[2]
astart=0.1
aslut=21
for (praecision in seqp)
{
for (j in seqj)
{
par[1]=astart+(j-1)*(aslut-astart)/20
L[j]=champLike(cost1home,par)
}
jmax=which.max(L)
astart=max(0.0001,astart+(jmax-2)*(aslut-astart)/20)
aslut=astart+jmax*(aslut-astart)/20
}
avalg[i]=astart+(jmax-1)*(aslut-astart)/20
Lvalg[i]=L[jmax]
}
iopt=which.max(Lvalg)
par[1]=avalg[iopt]
par[3]=cvalg[iopt]*par[2]
return(par)
}

par.Ch.home<-champPar(cost1home)

AIChome12<-2*3-2*champLike(cost1home,par.Ch.home)
AIChome12

BIChome12<-3*log(n)-2*champLike(cost1home,par.Ch.home)
BIChome12

sim.Champ.cost1_home <- qchamp(sim.data, par.Ch.home)
sim.L_Champ=sim.Champ.cost1_home-0.001

VaRChamp<-quantile(sim.L_Champ, alfa)
VaRChamp

#TVaR
mean(sim.L_Champ[sim.L_Champ > quantile(sim.L_Champ, alfa[1])])
mean(sim.L_Champ[sim.L_Champ > quantile(sim.L_Champ, alfa[2])])
mean(sim.L_Champ[sim.L_Champ > quantile(sim.L_Champ, alfa[3])])
cbind(mean(sim.L_Champ[sim.L_Champ > quantile(sim.L_Champ, alfa[1])]),
mean(sim.L_Champ[sim.L_Champ > quantile(sim.L_Champ, alfa[2])]),
mean(sim.L_Champ[sim.L_Champ > quantile(sim.L_Champ, alfa[3])]))

```



## IV. Código R: Cálculo Probabilidades de los Perfiles

```
betamd<-c(0.969531,-0.140253,0.030160,-0.743127,0.020086,-0.031336,-0.000466,0)
betabi<-c(0.081659,-0.107191,0.118761,-1.002263,-0.048717,0.049139,-0.032154,0)
betahome<-c(1.019694,0.013758,0.145599,0.026245,0.013955,-0.040872,0)
matrizcor<-matrix(c(1,0.796210,0.0502072,0.796210,1,0.053115,0.0502072,0.053115,1),3,3)
matrizcor
mean<-0
#tengo que poner la exposición en el vector de beta y x?

#3 riesgos
#Perfil 1
x<-c(1,0,0.37,0.18,1,1,0,0)#Perfil 1
xh<-c(1,0,0.37,1,1,0,0)
xbetamd<-betamd%*%x
xbetabi<-betabi%*%x
xbetahome<-betahome%*%xh

xbetabi<-as.numeric(xbetabi)
xbeta<-matrix(c(xbetamd,xbetabi,xbetahome))
xbeta<-as.numeric(xbeta)
xbeta

pmvnorm(lower=-Inf, upper=xbeta,mean=mean,corr=matrizcor)

#2riesgos (excepto bi)
#Perfil 1

xbeta.exbi_upper<-matrix(c(xbetamd,+Inf,xbetahome))
xbeta.exbi_upper<-as.numeric(xbeta.exbi_upper)
xbeta.exbi_upper

xbeta.exbi_lower<-matrix(c(-Inf,xbetabi,-Inf))
xbeta.exbi_lower<-as.numeric(xbeta.exbi_lower)
xbeta.exbi_lower

pmvnorm(lower=xbeta.exbi_lower, upper=xbeta.exbi_upper,mean=mean,corr=matrizcor)

#2riesgos (excepto md)
#Perfil 1

xbeta.exmd_upper<-matrix(c(+Inf,xbetabi,xbetahome))
xbeta.exmd_upper<-as.numeric(xbeta.exmd_upper)
xbeta.exmd_upper

xbeta.exmd_lower<-matrix(c(xbetamd,-Inf,-Inf))
xbeta.exmd_lower<-as.numeric(xbeta.exmd_lower)
xbeta.exmd_lower

pmvnorm(lower=xbeta.exmd_lower, upper=xbeta.exmd_upper,mean=mean,corr=matrizcor)

#2riesgos (excepto home)
#Perfil 1

xbeta.exh_upper<-matrix(c(xbetamd,xbetabi,+Inf))
```

```

xbeta.exh_upper<-as.numeric(xbeta.exh_upper)
xbeta.exh_upper

xbeta.exh_lower<-matrix(c(-Inf,-Inf,xbetahome))
xbeta.exh_lower<-as.numeric(xbeta.exh_lower)
xbeta.exh_lower

pmvnorm(lower=xbeta.exh_lower, upper=xbeta.exh_upper,mean=mean,corr=matrizcor)

#1riesgo(md)
#Perfil 1

xbeta.md_upper<-matrix(c(xbetamd,+Inf,+Inf))
xbeta.md_upper<-as.numeric(xbeta.md_upper)
xbeta.md_upper

xbeta.md_lower<-matrix(c(-Inf,xbetabi,xbetahome))
xbeta.md_lower<-as.numeric(xbeta.md_lower)
xbeta.md_lower

pmvnorm(lower=xbeta.md_lower, upper=xbeta.md_upper,mean=mean,corr=matrizcor)

#1riesgos (bi)
#Perfil 1

xbeta.bi_upper<-matrix(c(+Inf,xbetabi,+Inf))
xbeta.bi_upper<-as.numeric(xbeta.bi_upper)
xbeta.bi_upper

xbeta.bi_lower<-matrix(c(xbetamd,-Inf,xbetahome))
xbeta.bi_lower<-as.numeric(xbeta.bi_lower)
xbeta.bi_lower

pmvnorm(lower=xbeta.bi_lower, upper=xbeta.bi_upper,mean=mean,corr=matrizcor)

#1riesgos (home)
#Perfil 1

xbeta.h_upper<-matrix(c(+Inf,+Inf,xbetahome))
xbeta.h_upper<-as.numeric(xbeta.h_upper)
xbeta.h_upper

xbeta.h_lower<-matrix(c(xbetamd,xbetabi,-Inf))
xbeta.h_lower<-as.numeric(xbeta.h_lower)
xbeta.h_lower

pmvnorm(lower=xbeta.h_lower, upper=xbeta.h_upper,mean=mean,corr=matrizcor)

#Perfil 2
x<-c(1,0,0.19,0.1,0,0,0,1)#Perfil 2
xh<-c(1,0,0.19,0,0,0,1)
xbetamd<-betamd%*%x
xbetabi<-betabi%*%x
xbetahome<-betahome%*%xh

xbetabi<-as.numeric(xbetabi)
xbeta<-matrix(c(xbetamd,xbetabi,xbetahome))

```

```

xbeta<-as.numeric(xbeta)
xbeta

pmvnorm(lower=-Inf, upper=xbeta,mean=mean,corr=matrizcor)

#2riesgos (excepto bi)
#Perfil 2

xbeta.exbi_upper<-matrix(c(xbetamd,+Inf,xbetahome))
xbeta.exbi_upper<-as.numeric(xbeta.exbi_upper)
xbeta.exbi_upper

xbeta.exbi_lower<-matrix(c(-Inf,xbetabi,-Inf))
xbeta.exbi_lower<-as.numeric(xbeta.exbi_lower)
xbeta.exbi_lower

pmvnorm(lower=xbeta.exbi_lower, upper=xbeta.exbi_upper,mean=mean,corr=matrizcor)

#2riesgos (excepto md)
#Perfil 2

xbeta.exmd_upper<-matrix(c(+Inf,xbetabi,xbetahome))
xbeta.exmd_upper<-as.numeric(xbeta.exmd_upper)
xbeta.exmd_upper

xbeta.exmd_lower<-matrix(c(xbetamd,-Inf,-Inf))
xbeta.exmd_lower<-as.numeric(xbeta.exmd_lower)
xbeta.exmd_lower

pmvnorm(lower=xbeta.exmd_lower, upper=xbeta.exmd_upper,mean=mean,corr=matrizcor)

#2riesgos (excepto home)
#Perfil 2

xbeta.exh_upper<-matrix(c(xbetamd,xbetabi,+Inf))
xbeta.exh_upper<-as.numeric(xbeta.exh_upper)
xbeta.exh_upper

xbeta.exh_lower<-matrix(c(-Inf,-Inf,xbetahome))
xbeta.exh_lower<-as.numeric(xbeta.exh_lower)
xbeta.exh_lower

pmvnorm(lower=xbeta.exh_lower, upper=xbeta.exh_upper,mean=mean,corr=matrizcor)

#1riesgo(md)
#Perfil 2

xbeta.md_upper<-matrix(c(xbetamd,+Inf,+Inf))
xbeta.md_upper<-as.numeric(xbeta.md_upper)
xbeta.md_upper

xbeta.md_lower<-matrix(c(-Inf,xbetabi,xbetahome))
xbeta.md_lower<-as.numeric(xbeta.md_lower)
xbeta.md_lower

pmvnorm(lower=xbeta.md_lower, upper=xbeta.md_upper,mean=mean,corr=matrizcor)

```

```

#1riesgos (bi)
#Perfil 2

xbeta.bi_upper<-matrix(c(+Inf,xbetabi,+Inf))
xbeta.bi_upper<-as.numeric(xbeta.bi_upper)
xbeta.bi_upper

xbeta.bi_lower<-matrix(c(xbetamd,-Inf,xbetahome))
xbeta.bi_lower<-as.numeric(xbeta.bi_lower)
xbeta.bi_lower

pmvnorm(lower=xbeta.bi_lower, upper=xbeta.bi_upper,mean=mean,corr=matrizcor)

#1riesgos (home)
#Perfil 2

xbeta.h_upper<-matrix(c(+Inf,+Inf,xbetahome))
xbeta.h_upper<-as.numeric(xbeta.h_upper)
xbeta.h_upper

xbeta.h_lower<-matrix(c(xbetamd,xbetabi,-Inf))
xbeta.h_lower<-as.numeric(xbeta.h_lower)
xbeta.h_lower

pmvnorm(lower=xbeta.h_lower, upper=xbeta.h_upper,mean=mean,corr=matrizcor)

#Perfil 3
x<-c(1,1,0.25,0.03,3,0,1,0)#Perfil 3
xh<-c(1,1,0.25,3,0,1,0)
xbetamd<-betamd%*%x
xbetabi<-betabi%*%x
xbetahome<-betahome%*%xh

xbetabi<-as.numeric(xbetabi)
xbeta<-matrix(c(xbetamd,xbetabi,xbetahome))
xbeta<-as.numeric(xbeta)
xbeta

pmvnorm(lower=-Inf, upper=xbeta,mean=mean,corr=matrizcor)

#2riesgos (excepto bi)
#Perfil 3

xbeta.exbi_upper<-matrix(c(xbetamd,+Inf,xbetahome))
xbeta.exbi_upper<-as.numeric(xbeta.exbi_upper)
xbeta.exbi_upper

xbeta.exbi_lower<-matrix(c(-Inf,xbetabi,-Inf))
xbeta.exbi_lower<-as.numeric(xbeta.exbi_lower)
xbeta.exbi_lower

pmvnorm(lower=xbeta.exbi_lower, upper=xbeta.exbi_upper,mean=mean,corr=matrizcor)

#2riesgos (excepto md)
#Perfil 3

```

```

xbeta.exmd_upper<-matrix(c(+Inf,xbetabi,xbetahome))
xbeta.exmd_upper<-as.numeric(xbeta.exmd_upper)
xbeta.exmd_upper

xbeta.exmd_lower<-matrix(c(xbetamd,-Inf,-Inf))
xbeta.exmd_lower<-as.numeric(xbeta.exmd_lower)
xbeta.exmd_lower

pmvnorm(lower=xbeta.exmd_lower, upper=xbeta.exmd_upper,mean=mean,corr=matrizcor)

#2riesgos (excepto home)
#Perfil 3

xbeta.exh_upper<-matrix(c(xbetamd,xbetabi,+Inf))
xbeta.exh_upper<-as.numeric(xbeta.exh_upper)
xbeta.exh_upper

xbeta.exh_lower<-matrix(c(-Inf,-Inf,xbetahome))
xbeta.exh_lower<-as.numeric(xbeta.exh_lower)
xbeta.exh_lower

pmvnorm(lower=xbeta.exh_lower, upper=xbeta.exh_upper,mean=mean,corr=matrizcor)

#1riesgo(md)
#Perfil 3

xbeta.md_upper<-matrix(c(xbetamd,+Inf,+Inf))
xbeta.md_upper<-as.numeric(xbeta.md_upper)
xbeta.md_upper

xbeta.md_lower<-matrix(c(-Inf,xbetabi,xbetahome))
xbeta.md_lower<-as.numeric(xbeta.md_lower)
xbeta.md_lower

pmvnorm(lower=xbeta.md_lower, upper=xbeta.md_upper,mean=mean,corr=matrizcor)

#1riesgos (bi)
#Perfil 3

xbeta.bi_upper<-matrix(c(+Inf,xbetabi,+Inf))
xbeta.bi_upper<-as.numeric(xbeta.bi_upper)
xbeta.bi_upper

xbeta.bi_lower<-matrix(c(xbetamd,-Inf,xbetahome))
xbeta.bi_lower<-as.numeric(xbeta.bi_lower)
xbeta.bi_lower

pmvnorm(lower=xbeta.bi_lower, upper=xbeta.bi_upper,mean=mean,corr=matrizcor)

#1riesgos (home)
#Perfil 3

xbeta.h_upper<-matrix(c(+Inf,+Inf,xbetahome))
xbeta.h_upper<-as.numeric(xbeta.h_upper)
xbeta.h_upper

```

```

xbeta.h_lower<-matrix(c(xbetamd,xbetabi,-Inf))
xbeta.h_lower<-as.numeric(xbeta.h_lower)
xbeta.h_lower

pmvnorm(lower=xbeta.h_lower, upper=xbeta.h_upper,mean=mean,corr=matrizcor)

#Perfil 4
x<-c(1,1,0.60,0.41,1,1,0,0)#Perfil
xh<-c(1,1,0.60,1,1,0,0)
xbetamd<-betamd%*%x
xbetabi<-betabi%*%x
xbetahome<-betahome%*%xh

xbetabi<-as.numeric(xbetabi)
xbeta<-matrix(c(xbetamd,xbetabi,xbetahome))
xbeta<-as.numeric(xbeta)
xbeta

pmvnorm(lower=-Inf, upper=xbeta,mean=mean,corr=matrizcor)

#2riesgos (excepto bi)
#Perfil 4

xbeta.exbi_upper<-matrix(c(xbetamd,+Inf,xbetahome))
xbeta.exbi_upper<-as.numeric(xbeta.exbi_upper)
xbeta.exbi_upper

xbeta.exbi_lower<-matrix(c(-Inf,xbetabi,-Inf))
xbeta.exbi_lower<-as.numeric(xbeta.exbi_lower)
xbeta.exbi_lower

pmvnorm(lower=xbeta.exbi_lower, upper=xbeta.exbi_upper,mean=mean,corr=matrizcor)

#2riesgos (excepto md)
#Perfil 4

xbeta.exmd_upper<-matrix(c(+Inf,xbetabi,xbetahome))
xbeta.exmd_upper<-as.numeric(xbeta.exmd_upper)
xbeta.exmd_upper

xbeta.exmd_lower<-matrix(c(xbetamd,-Inf,-Inf))
xbeta.exmd_lower<-as.numeric(xbeta.exmd_lower)
xbeta.exmd_lower

pmvnorm(lower=xbeta.exmd_lower, upper=xbeta.exmd_upper,mean=mean,corr=matrizcor)

#2riesgos (excepto home)
#Perfil 4

xbeta.exh_upper<-matrix(c(xbetamd,xbetabi,+Inf))
xbeta.exh_upper<-as.numeric(xbeta.exh_upper)
xbeta.exh_upper

xbeta.exh_lower<-matrix(c(-Inf,-Inf,xbetahome))
xbeta.exh_lower<-as.numeric(xbeta.exh_lower)
xbeta.exh_lower

```

```

pmvnorm(lower=xbeta.exh_lower, upper=xbeta.exh_upper,mean=mean,corr=matrizcor)

#1riesgo(md)
#Perfil 4

xbeta.md_upper<-matrix(c(xbetamd,+Inf,+Inf))
xbeta.md_upper<-as.numeric(xbeta.md_upper)
xbeta.md_upper

xbeta.md_lower<-matrix(c(-Inf,xbetabi,xbetahome))
xbeta.md_lower<-as.numeric(xbeta.md_lower)
xbeta.md_lower

pmvnorm(lower=xbeta.md_lower, upper=xbeta.md_upper,mean=mean,corr=matrizcor)

#1riesgos (bi)
#Perfil 4

xbeta.bi_upper<-matrix(c(+Inf,xbetabi,+Inf))
xbeta.bi_upper<-as.numeric(xbeta.bi_upper)
xbeta.bi_upper

xbeta.bi_lower<-matrix(c(xbetamd,-Inf,xbetahome))
xbeta.bi_lower<-as.numeric(xbeta.bi_lower)
xbeta.bi_lower

pmvnorm(lower=xbeta.bi_lower, upper=xbeta.bi_upper,mean=mean,corr=matrizcor)

#1riesgos (home)
#Perfil 4

xbeta.h_upper<-matrix(c(+Inf,+Inf,xbetahome))
xbeta.h_upper<-as.numeric(xbeta.h_upper)
xbeta.h_upper

xbeta.h_lower<-matrix(c(xbetamd,xbetabi,-Inf))
xbeta.h_lower<-as.numeric(xbeta.h_lower)
xbeta.h_lower

pmvnorm(lower=xbeta.h_lower, upper=xbeta.h_upper,mean=mean,corr=matrizcor)

```

## Referencias

- [1] Amemiya, T., 1985b. *Advanced Econometrics*. Harvard University Press, Cambridge, Massachusetts.
- [2] Artzner, P., Delbaen, F., Eber, J. M. y Heath, D. (1999a). Coherent measures of risk. *Mathematical Finance*.
- [3] Bolancé, C., Guillén, M., Padilla, A., 2015a. Estimación del riesgo mediante el ajuste de cópulas. *UB Riskcenter Working Paper Series*.
- [4] Bolancé, C., Guillén, M., Gustafsson, J., Nielsen, J., 2012b. *Quantitative Operational Risk Models*. Chapman Hall/CRC Finance Series, London.
- [5] Buch-Larsen, T., Perch Nielsen, J., Guillén, M., Bolancé, C., 2005a. Kernel density estimation for heavy-tailed distributions using the champernowne transformation. *Statistics* Vol. 39, No. 6, 503–518.
- [6] Cayuela, L., 2010a. Modelos lineales generalizados (GLM). *Materiales de un curso del R del IREC*.
- [7] Chib, S, Greenberg, E., 1998a, Analysis of multivariate probit models. *Biometrika*.
- [8] De Jong, P., Heller G.Z., 2008b. *Generalized Linear Models for Insurance Data*. Cambridge University Press Cambridge.
- [9] Frees, E. W. 2010b, *Regression Modeling with Actuarial and Financial Applications*. Cambridge University Press Cambridge.
- [10] Frees, E. W., Valdez, E. A., 1998a, Understanding Relationships Using Copulas. *North American Actuarial Journal*. Vol. 2, No. 1, 1–25.
- [11] Greene, W.H., 1983b. *Econometric Analysis*. Nueva Jersey.



- [12] McCullagh, P., Nelder, J.A., 1983b. *Generalized Linear Models*. Chapman and Hall, London.
- [13] Moscote Flórez, O., Rincón, W.A., 2012a. Modelo Logit y Probit : un caso de aplicación. *Universidad Santo Tomás*.
- [14] Schmidheiny, K., 2016a. Binary Response Models. Short Guides to Microeconometrics, *Universität Basel*.
- [15] Torres, G. I., Olarte, A. M., 2009a. Valor en riesgo desde un enfoque de cópulas. *Universidad EAFIT Medellín*.
- [16] Tse, Y. K., 2009b. *Nonlife Actuarial Models*. Cambridge University Press.
- [17] Young, G., Valdez, E. A., Kohn, R., 2006a. Multivariate probit models for conditional claim-types. *Insurance: Mathematics and Economics*. Vol. 44, No. 2, 214-228.