uc3m | Universidad Carlos III de Madrid

Máster Universitario en Ciencias Actuariales y Financieras
2020-2021

Trabajo Fin de Máster

"Modelización de la prima de ciberseguro con funciones cópula"

Antonio Yaque Aguilera

Tutor/es

José Miguel Rodríguez-Pardo del Castillo

Jesús Ramón Simón del Potro

Madrid, 15 de junio de 2021

RESUMEN

El contrato de ciberseguro supone una variable de interés desde el punto de vista financiero-actuarial al representar una herramienta fundamental para la gestión de riesgos, ya que ayudará a controlarlos y mitigarlos. El presente trabajo tiene como objetivo proponer una metodología, basada en las funciones cópula, que pueda usarse para cuantificar la prima anual que repercutirán las aseguradoras de cara a aminorar los impactos adversos de estos riesgos, que se materializan en ciberataques y errores humanos, entre otros. Para introducir conceptos como "ciberataque" se abarcará teóricamente el ciberseguro en sí y este tipo de riesgo, para lograr una mayor comprensión de la situación actual en dicho contexto. La investigación llevada a cabo permite tarificar diferentes pólizas de ciberseguro, con lo cual se observará el efecto de incluir diferentes elementos del reaseguro.

Palabras clave: seguro, ciberriesgo, gestión de riesgos, prima, cópula, tarificación

ABSTRACT

The cyberinsurance contract is a variable of interest from the financial-actuarial point of view, as it represents a fundamental tool for risk management, since it will help control and mitigate risks. The aim of this paper is to propose a methodology, based on copula functions, that can be used to quantify the annual premium that insurers will charge to mitigate the adverse impacts of these risks, which are materialized in cyber-attacks and human errors, among others. In order to introduce concepts such as "cyber-attack", cyber-insurance itself and this type of risk will be covered theoretically, so as to achieve a better understanding of the current situation in this context. The research carried out allows the pricing of different cyber insurance policies, thus the effect of including different elements of reinsurance will be observed.

Keywords: insurance, cyber-risk, risk management, premium, copula, pricing

DEDICATORIA

A mis dos ángeles, por todo el tiempo que pasamos juntos aunque no os pudiera ver; “ni nada ni nadie”. A mi madre, mis hermanos, mi tía y mi abuelo. Yo hoy no estaría aquí sin ustedes.

ÍNDICE DE CONTENIDOS

ÍNDICE DE FIGURAS

ÍNDICE DE TABLAS

1. INTRODUCCIÓN

Hoy día es imposible pensar en nuestra vida sin Internet, ya que actualmente todo está conectado a una red. Así pues, gracias a él hacemos la vida tal y como la conocemos y gran parte del desarrollo tecnológico del último siglo ha sido llevado a cabo gracias al mismo. Sin embargo, a pesar de usarlo de forma habitual en nuestro día a día no somos plenamente conscientes de los muchos riesgos que entraña, los cuales se dan cada vez con más frecuencia y provocan grandes problemas para ciudadanos y empresas por igual.

Como se verá, la gestión y la cobertura de lo que se denominará en adelante como ciberriesgo, o riesgo cibernético, se han convertido en un proceso esencial para instituciones públicas, pymes, entidades aseguradoras y grandes empresas en su conjunto. Estas organizaciones interactúan constantemente con un ambiente, más bien impreciso, en el que necesitan una evaluación aproximada de dichos riesgos para controlarlos o mitigarlos, proporcionando así un servicio adecuado y de calidad a sus clientes. En este sentido, entre las herramientas a su alcance se encuentra la contratación de una póliza de ciberseguro, para cuya tarificación se propondrá una metodología basada en el uso de cópulas. Es importante destacar que este sector "moverá este año en España un volumen de negocio cercano a los 1.320 millones de euros, lo que supone un 8,1% más que en 2020" (Servimedia, 2021).

1.1. OBJETIVOS

En primer lugar, algo que debe ser tenido en consideración es el horizonte temporal escogido para el cumplimiento de los objetivos del presente estudio, situándonos en este caso en dos ventanas temporales, 2003 y 2020. Con ello, el propósito del trabajo será proponer una metodología, basada en simulación mediante cópulas, para estimar las primas anuales de ciberseguro en estos dos periodos a partir de los datos reportados por ICSA (Bridwell, 2004; en Hearth y Herath, 2011) y NetDiligence (2020), compañías encargadas de valorar el ciberriesgo diferenciando los principales ciberriesgos a los que se enfrentan las empresas, para lo que reportan el número de ordenadores afectados y el importe total de la pérdida por el incidente.

De esta manera se realizará un contraste entre la situación dada en 2003, cuando el uso de Internet no era común, y la que se da hoy día. Además, se considerarán tres tipos diferentes de pólizas para ver cómo influyen los elementos típicos del reaseguro en la modelización de las primas.

Una vez conseguido esto, las aseguradoras podrán usar las estimaciones para comenzar los procedimientos habituales que sirven para protegerse frente a posibles impactos derivados de pérdidas no esperadas, incluyendo otros importes adicionales para llegar a su prima comercial. De esta forma, esta metodología contribuiría a la literatura en cuanto a la tarificación del ciberseguro al tratar un enfoque actuarial que permita, entre otras cosas, tener en cuenta elementos fundamentales del contrato de seguro para mitigar el riesgo en un contexto de amenazas crecientes como resultado del rápido desarrollo tecnológico.

En este sentido David Garrido, director de siniestros de AGCS España, explica que:

"La exposición a los ciberataques ha ido creciendo y el año pasado tuvimos un número récord de incidentes. El impacto de los siniestros cibernéticos reportados (…) en 2020 fue tres veces mayor que el año anterior. Asimismo, los gerentes de riesgos son cada vez más conscientes de la necesidad de proteger sus sistemas y datos, buscando por ello una protección integral de las aseguradoras que operan en este mercado" (Leonor, 2021).

De ahí la primera razón de la importancia de los ciberseguros: una decisión de inversión que no contemple adquirir un seguro de este tipo desembocaría en un decremento de la protección de la empresa y, por ende, en un incremento de las pérdidas asociadas a los riesgos provenientes del entorno cibernético. Así Santiago Gutiérrez, socio líder de riesgos cibernéticos en Deloitte México, reporta que:

"Lo fundamental es prevenir y protegerse (…) de las situaciones adversas que hay en materia cibernética y que se seguirán presentando. Las consecuencias de los ciberataques se traducen en gastos no presupuestados por las organizaciones, ya que reaccionar ante un incidente de esa naturaleza implica contratar a terceros o destinar recursos extra para recuperar la operación del negocio" (Deloitte, 2019).

Con todo, el principal objetivo de este Trabajo Fin de Máster será modelizar, mediante el uso de cópulas, la prima anual que permitiría a las empresas cubrirse del ciberriesgo a través de la herramienta estadística R, la cual está al alcance de todos. Para ello, se establecerá como objetivo secundario el realizar un análisis teórico del ciberriesgo, aspecto esencial para comprender el funcionamiento del seguro en sí: entendiendo los principales ingredientes que lo componen podrá elaborarse el caso de estudio práctico.

1.2. ESTRUCTURA

La estructura que se ha seguido en la elaboración de este trabajo ha sido gradual, empezando desde lo más general para, al final, acabar con datos precisos y específicos. Constará de seis capítulos, además de este apartado introductorio.

En el capítulo primero se ha realizado una breve introducción acerca de la importancia del seguro de ciberriesgos en el ámbito empresarial, así como un desglose de los objetivos perseguidos en este trabajo y de su estructura.

Por su parte, el capítulo dos hace una breve reseña sobre las causas y las consecuencias que ha tenido el ciberriesgo desde sus inicios hasta la actualidad, resaltando así la amenaza creciente para todos los agentes de la economía moderna. Además, se describirán los principales ciberataques, agrupándolos en categorías con el fin de que el lector se familiarice con la terminología propuesta. A la vez, se realizará un análisis de la situación en los últimos años en España con el fin de ver hasta qué punto se abarca este problema y qué soluciones se han ido proponiendo en consecuencia.

En el tercer capítulo se hará una introducción a los ciberseguros, los cuales representan una vía para reducir o mitigar dichos riesgos. De esta forma se repasarán las principales coberturas que ofrecen, así como los principales clientes que se decantan por estos contratos por la vulnerabilidad y la exposición que poseen. Por otra parte, se harán varias conclusiones sobre la importancia del reaseguro en el sector, incluyendo los principales contratos que pueden encontrarse: proporcionales y no proporcionales con el fin de ver los elementos principales que intervendrán en la modelización del ciberseguro.

En el cuarto capítulo del estudio se procederá a modelizar posibles escenarios de propagación de los virus con los modelos epidemiológicos SIS y SIR, para posteriormente hacer lo propio con las primas de ciberseguro con datos de 2003 y 2020, usando para ello las funciones cópula más adecuadas y viendo qué distribuciones marginales se ajustan mejor a las variables estudiadas. Una vez obtenido esto, se simulará una muestra bivariante con la que se obtendrá la función de distribución de pérdidas por el pago de la suma asegurada, elemento necesario como parte de la ecuación que devuelve el precio de la prima de seguro. Una vez que esto se consiga se propondrán varios tipos de pólizas con diferentes elementos para observar el efecto del reaseguro sobre las primas.

Posteriormente, en el capítulo cinco se expondrán los resultados obtenidos, aspecto fundamental del trabajo. Además, el sexto recogerá las referencias bibliográficas que han sido necesarias para redactar el presente documento.

Por último, y a modo de anexo, se expone un apartado que presenta el código programado en el *software* R, aplicado para la obtención de los resultados.

2. EL CIBERRIESGO, CAUSAS Y CONSECUENCIAS

En primer lugar cabe introducir el concepto de ciberriesgo. El Manual de Tallín (CCDCOE, 2012) es un informe creado para difundir el protocolo a seguir bajo el supuesto de aparición de una ciberguerra, definiendo el ciberriesgo como "toda operación cibernética, ya sea ofensiva o defensiva, que se espera razonablemente que cause lesiones o la muerte a personas o daños o destrucción de objetos" (CCDCOE, 2012).

Tradicionalmente se ha considerado que este tipo de riesgo ha sido objeto exclusivo de estudio de la informática, estando el resto exentos de concienciación y preocupación respecto a ellos. Sin embargo, los peligros que genera el tráfico de datos en la web es uno de los riesgos más importantes para las entidades, en general, y para los individuos en particular, con lo que la sociedad en su conjunto debería estar al tanto de la tendencia que puedan seguir estos riesgos, profundizar en ellos y buscar una manera de poder combatirlos. En el contexto actuarial, la respuesta se encuentra en los seguros sobre ciberriesgos, o ciberseguros.

2.1. BREVE RESEÑA HISTÓRICA SOBRE LA CIBERSEGURIDAD

Con la aparición de Internet en 1960 a través de ARPANET se descubre un modo de vida completamente diferente del que existía y con él se abrió la puerta a nuevas amenazas hasta entonces desconocidas, que debían y deben ser controladas y gestionadas en la medida de lo posible. Así, se creó una situación de vulnerabilidad progresiva en lo referente a ellas debido, sobre todo, a la creciente dependencia que existe en el ámbito empresarial y personal con los sistemas informáticos.

Los inicios de la ciberseguridad se remontan a los años 70 con el surgimiento del primer virus cibernético conocido, Creeper, el cual puso de evidencia la necesidad de un sistema que protegiera la seguridad de los datos (López, 2019). Así fue como en 1972 nació el antivirus Reaper, creado específicamente para combatirlo; "decimos que no es un antivirus en sí, ya que en realidad era un virus estilo gusano porque se autorreplicaba y se extendía a través de la red, pero teniendo características similares a un antivirus" (Wikipedia, 2021).

Como sucede con todos los virus estos comenzaron a propagarse, en tanto que la tecnología fue desarrollándose cada vez más rápido, lo que conllevó a un mayor volumen de datos en la red y a un aumento de los riesgos asociados (López, 2019). Al haber tantos estos escaparon del control de los antivirus, con lo que entró en juego la Endpoint Detection and Response (EDP) que son "recursos para combatir las amenazas avanzadas y responder a incidentes en los puntos finales de la red (…) [y] proporcionan detalles forenses que permiten ofrecer una respuesta rápida ante incidentes" (Arsys, 2019).

En este sentido, el siguiente gran reto que debía ser afrentado sería el surgimiento del Internet of Things (IoT), con mayor presencia aún si cabe en nuestros días, puesto que los dispositivos *wearable*, de uso cotidiano, son también susceptibles de ataques cibernéticos. De esta forma, son varios los aspectos que deben considerarse (ValoraData, 2018):

i) Incorporación de dispositivos al sistema IoT: a través de la programación es posible incorporar mecanismos para facilitar controles de seguridad de forma automática. Esto sería lo ideal para sistemas cerrados como aplicaciones integradas en los diferentes dispositivos.

ii) Seguridad controlada por los propietarios: para sistemas abiertos a través de la web donde es el usuario el que debe controlar la propia seguridad.

iii) Autenticación IoT: uso de credenciales para poder acceder a determinadas secciones de las apps.

iv) Privacidad e integridad de los datos: más que proteger los dispositivos en sí, de lo que se trata es de proteger los datos que circulan a través de ellos. Así, estos deberán estar cifrados para alcanzar un mayor grado de protección.

v) Actualizaciones de *Firmware* seguras.

2.2. CIBERATAQUES: CONCEPTO Y CLASIFICACIÓN

Un buen comienzo para llegar a concienciarnos en profundidad de los dilemas que conllevan los ciberataques es comprender su significado y saber sus variantes. Así, la RAE (2021) define el ciberataque como un "ataque organizado contra el sistema informático de una entidad o empresa con el objetivo de bloquearlo, dañarlo u obtener información", pudiendo ser usados incluso como "armas" en ciberguerras o ataques terroristas. Para poder introducir al lector a la terminología que se usará en lo sucesivo, en este apartado se definirán los más empleados por los hackers.

2.2.1. Fugas de información

Este tipo de ciberataque se da cuando:

"Algún dato o activo de información que tenga valor para una organización pasa a manos ajenas, perdiendo la cualidad de confidencialidad que le fue asignada. Esto se puede ver representado, por ejemplo, en documentos que pasan a ser accesibles por personas no autorizadas, o también por cualquier dato secreto que alguien interno le facilite a un externo sin pasar por un medio digital" (Pacheco, 2011).

Es bastante común encontrar casos de robos de información, sobre todo en la última década, siendo uno de los principales ejemplo el caso Wikileaks, organización que desde el 2006 autoriza a usuarios a publicar en sus páginas información confidencial de interés público ocultando su identidad, escudándose en el anonimato para hacerlo. Esto tuvo una gran repercusión a nivel mundial, ya que Estados Unidos sufrió "la mayor filtración de documentos secretos de la historia" (Pacheco, 2011).

2.2.2. *Malware*

Acrónimo de "*malicious software*". Es un programa informático que tiene como finalidad introducirse en el dispositivo y sustraer o dañar la información. Dentro del *malware* se distingue entre:

- Virus informático: conocido por todos. Una vez que el usuario lo abre infecta archivos y puede extenderse por el ordenador sin control, subrayando que debe ser el usuario el que clique en ellos. Así, puede ser evitado con la instalación de un antivirus.

- Gusano informático: de finalidad similar a la del virus, pero a diferencia de este no requiere que sea abierto por el usuario. Este tipo de *malware* hace que se ralenticen las funciones de los dispositivos, de ahí su nombre, por lo que sería algo más complejo de descubrir.

- Troyano: es un programa que, como en la leyenda griega, no presenta una apariencia maliciosa. Sin embargo, una vez dentro del ordenador permite a los hackers hacerse con el control de los dispositivos. A diferencia de los casos anteriores, este *malware* no se expande por sí mismo.

- *Spyware*: su función es "espiar" información acerca del usuario, como su historial de Internet, las acciones que realiza (mediante capturas de pantalla), los archivos que posee, entre otras cosas; todo ello de manera silenciosa y en un segundo plano para pasar desapercibido.

- *Adware*: este tipo de *software* malicioso no tiene otro objetivo que mostrar de manera repetida e incesante anuncios publicitarios, con lo que el dispositivo queda infectado por pantallas emergentes publicitarias. Hay quien considera que también permite que los equipos puedan ser monitorizados por los ciberdelicuentes.

- *Ransomware*: impide de manera taxativa a los usuarios abrir el archivo infectado, ya que para ello pide un rescate que deberá ser pagado para eliminar el cifrado que crean sobre ellos.

2.2.3. *Phishing*

"Pesca de datos", es la forma más antigua de ciberataque. Este ataque consiste en hacer creer al usuario, a través generalmente de correos electrónicos infectados, que debe hacer clic en un enlace malicioso que adjuntan haciéndose pasar por una entidad bancaria o cualquier empresa en concreto. Es así como intentan conseguir datos personales como contraseñas, tarjetas de crédito o demás datos privados para llevar a cabo una estafa.

De esta manera, como se observa en la *Tabla 2.1*, Gendre (2020) reporta en VadeSecure que Microsoft es la principal marca suplantada en ataques de phishing, seguida por Facebook, PayPal y eBay, entre otros:

1.	Microsoft
2.	Facebook
3.	PayPal
4.	eBay
5.	Chase
6.	Amazon
7.	Netflix
8.	WhatsApp
9.	DHL
10.	Google

Tabla 2.1 Principales marcas suplantadas por "*phishers*"

Fuente: Elaboración propia con datos de VadeSecure (Gendre, 2020)

Además, a partir de febrero de 2020, coincidiendo con la pandemia del COVID-19 la compañía de *software* estadounidense Symantec reporta, como se expone en la *Figura 2.1*, un aumento considerable de bloqueos de emails *phishing* que llevaban de asunto palabras como "coronavirus", "corona" o "COVID-19", situación aprovechada por los ciberdelicuentes para llevar a cabo sus estafas; más aún aprovechando el auge del teletrabajo:

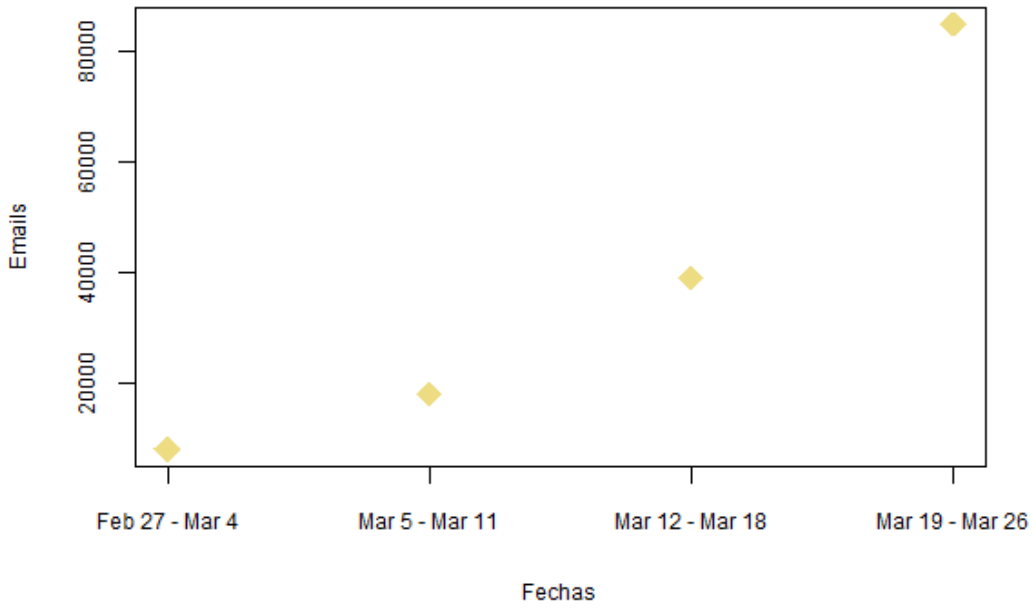

Figura 2.1 Emails phishing referentes al coronavirus

Fuente: Elaboración propia con datos de Symantec (Thaware, 2020)

2.2.4. *Formjacking*

Tiene como objetivo robar datos de tarjetas de crédito, al realizar compras *online*, para luego venderlos en el mercado negro. Como se reporta en un estudio realizado durante finales de verano: "si comparamos la semana del 13 al 20 de septiembre con la misma semana de agosto, la cantidad de casos de *formjacking* bloqueados por Symantec se duplicó con creces, pasando de poco más de 41.000 a casi 88.500, un aumento porcentual del 117%" (Symantec, 2018). Con esto, estos ciberataques fueron los que más se expandieron en 2018.

2.2.5. *Man-in-the-Middle-Attack*

También conocido como Ataque de Intermediario o MitM, donde existe un intermediario entre el hacker y el usuario, que puede ser la cuenta de correo o la bancaria, entre otros. De esta manera, el atacante podría introducirse en la red WiFi para interceptar los datos privados que circulan por estos canales.

2.2.6. *Distributed Denial of Service* (DDoS)

El objetivo principal de estos ataques es causar una denegación de servicios en los dispositivos afectados a través de un colapso de los mismos al "bombardear" de solicitudes su servidor, con lo cual tratan de sobrecargarlos, consiguiendo una denegación del servicio.

2.3. ANÁLISIS DE LA CIBERSEGURIDAD EN ESPAÑA

Tradicionalmente, España siempre ha sido uno de los países más propensos a recibir ciberataques. Así, de cara a ilustrar y dar a conocer en profundidad la dimensión del problema que se plantea, cabe mencionar lo que ocurrió en 2017 con el *ransomware* WannaCry, el ciberataque con más repercusión a nivel mundial. Este virus llegó a encriptar más de 200,000 ordenadores en más de 150 países, con los consiguientes colapsos en numerosas compañías de todos los sectores.

Como se puede observar en la *Tabla 2.2*, según Europol y la OTAN (2018) en España la compañía más afectada fue Telefónica, si bien otras como Iberdrola y Gas Natural también recibieron un gran impacto:

País	Empresa más afectada
Alemania	Deutsche Bahn 02
Eslovaquia	Faculty Hospital, Nitra
España	Telefónica
Hungría	Telenor Hungary
Portugal	Portugal Telecom
Reino Unido	National Health Service & Nissan
Rusia	Russian Railways & Ministry of Internal Affairs of the Russian Federation

Tabla 2.2 Objetivos destacados de WannaCry en la Unión Europea

Fuente: Elaboración propia con datos de Europol y OTAN (2018)

El modo de actuar de WannaCry era el siguiente: los ciberdelincuentes pedían un rescate en Bitcoin para recuperar los datos secuestrados. Si esta cantidad no era pagada pasadas 72 horas, el precio se duplicaría; y si en otras 72 horas seguía sin pagarse, entonces el archivo quedaría permanentemente cifrado. A pesar de la repercusión esta no se tradujo en grandes pérdidas monetarias, pero sí en un aumento de la concienciación en cuanto a las medidas preventivas para evitar este tipo de incidentes.

"Podríamos decir que WannaCry ha vuelto a advertir una vez más de la necesidad de emplear todas estas medidas y que la prevención es uno de los factores claves en la estrategia de ciberseguridad de las compañías. Este ataque masivo solamente ha demostrado ser una pequeña parte del arsenal que aún no ha visto la luz. Debemos estar preparados para poder hacer frente a este tipo de amenazas" (Deloitte, 2017).

Aun así, sería un error pensar que este tipo de ataques van dirigidos únicamente a grandes corporaciones, puesto que también afectan a ciudadanos y pymes por igual y "no solo por la debilidad tecnológica de estas sino también por el desconocimiento sobre las ciberamenazas" (Google y The Cocktail Analysis, 2019). De esta forma, según se desprende del Balance de Ciberseguridad elaborado por el Instituto Nacional de Ciberseguridad (INCIBE, 2017-2019), los incidentes que se han registrado en los últimos años podrían agruparse tal y como se recoge en la *Figura 2.2*:

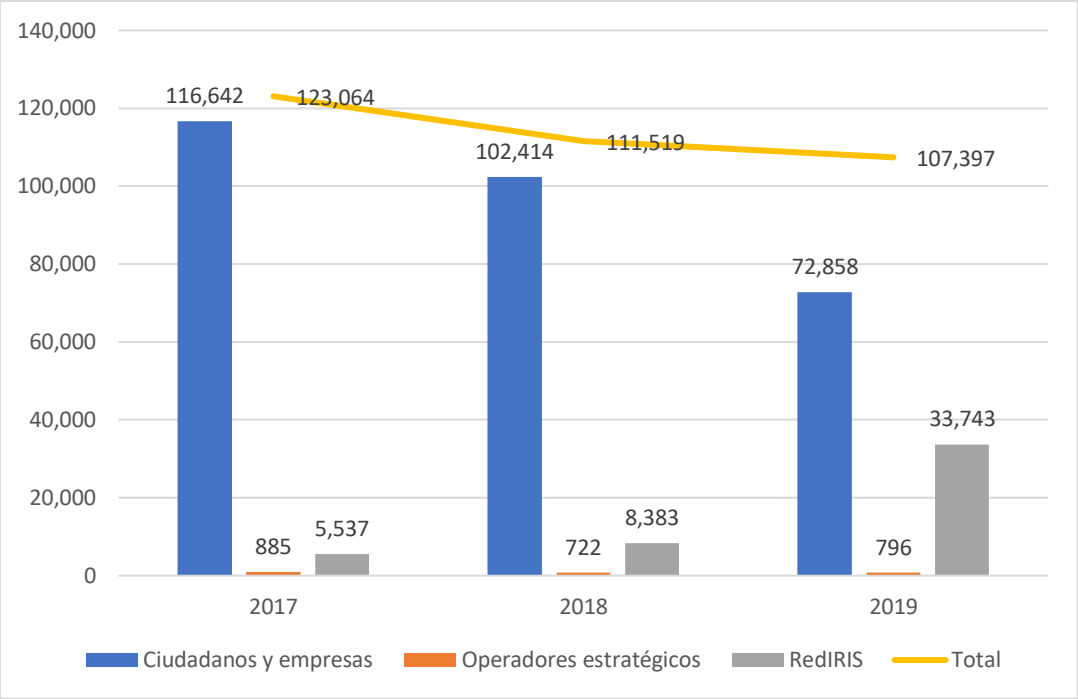

Figura 2.2 Incidentes gestionados por INCIBE

Fuente: Elaboración propia con datos de INCIBE (2017-2019)

Con ello, se observa que en los últimos años los ataques dirigidos a ciudadanos y empresas han sido de un 94.78%, un 91.83% y un 67.84% respectivamente, un porcentaje que va disminuyendo gracias a las medidas de concienciación, pero que no debe hacer que se baje la guardia frente a estas amenazas.

Una vez vistas estas dos grandes perspectivas cabe destacar que, en el plano de las grandes empresas, Vodafone (2018) consideró que España y otros países de la Unión Europea, como Italia, Alemania e Irlanda, adoptan una posición "reactiva" en materia de ciberseguridad. Sin embargo, países como Estados Unidos, Reino Unido o India son considerados en una posición "en desarrollo", lo que significaría que en nuestro país en las empresas "se ha tomado algunas medidas para asegurar su negocio, pero generalmente están en desventaja en lo que se refiere a la seguridad cibernética. Tienen un alcance significativo para mejora en todos los ámbitos" (Vodafone, 2018).

Desde el punto de vista de las pymes, a través de 720 encuestas se revela que, si bien estas están más concienciadas, esto no implica que adopten todas las medidas de prevención posibles: solo un 36% le da un alto grado de relevancia a la ciberseguridad al establecer protocolos básicos de seguridad, como la verificación de dos pasos en el correo de la empresa (Google y The Cocktail Analysis, 2019).

En este informe se recoge también que entre los usuarios, aunque también existe un mayor grado de concienciación esto no se traslada como se debería a la práctica: únicamente un 14% cambia sus contraseñas regularmente y un 21% hace copias de seguridad de sus archivos, además de actualizar los sistemas operativos de los dispositivos. Entre las principales conclusiones de su investigación sobre ciberseguridad en España se encuentran que, como se veía con lo sucedido a nivel mundial con WannaCry, la ciberseguridad "traspasa fronteras y requiere de una legislación transnacional" (Google y The Cocktail Analysis, 2019), lo cual se analizará en el siguiente apartado.

2.4. CONTEXTO LEGISLATIVO: AVANCES EN MATERIA DE REGULACIÓN

En primer lugar, entre los organismos que velan por la ciberseguridad se encuentran (IMF Business School, 2020):

a) Instituto Nacional de Ciberseguridad (INCIBE), que depende del Ministerio de Economía y Empresa y ofrece apoyo en cuanto a ciberseguridad a empresas, ciudadanos y universidades.

b) Centro Criptológico Nacional del Centro Nacional de Inteligencia (CCN-CERT), su "escisión tecnológica". Este CERT (siglas de *Computer Emergency Response Team*) es un organismo de carácter público que coordina las alertas y respuestas frente a problemas de ciberseguridad en el plano nacional. Así, su función principal será proteger los sistemas usados en la Administración Pública.

c) Centro Nacional para la Protección de las Infraestructuras Críticas (CNPIC): como su propio nombre indica, este órgano del Ministerio del Interior supervisa y coordina la protección de las infraestructuras críticas (proveedores de luz, agua, etc.) que están repartidas por toda España.

d) Mando Conjunto de Ciberdefensa (MCCD), que depende del Jefe de Estado Mayor de Defensa y por ende del Ministerio de Defensa. Este mecanismo permite ejercer movimientos de ciberdefensa ante ataques que pongan en jaque la ciberseguridad nacional.

Así, en este apartado se desarrollarán las principales líneas que anota este reporte encargado por Google y preparado por The Cocktail Analysis (2019), donde se indica que estos mecanismos deberán afrontar conjuntamente el reto que supone la ciberseguridad, la cual supone generalmente un desafío cuanto menos complejo dada la extensión de su dimensión.

De esta manera, se puede observar en la práctica que pocos ciberataques afectan exclusivamente a una región, dado que la mayoría van dirigidos a todos por igual. Además, es igualmente complicado intentar rastrear de dónde proceden los mismos, pues se amparan en el anonimato que proporcionan las redes (Google y The Cocktail Analysis, 2019). Estos serían los primeros obstáculos para la creación de un sistema legal que permita abordar y legislar sobre esta materia.

En este respecto David Barroso, fundador de la prestigiosa compañía española de ciberseguridad CounterCraft, afirma que "Internet es el salvaje oeste: es muy fácil de atacar y muy difícil de defender" (Google y The Cocktail Analysis, 2019). A pesar de ello, son muchos los avances en materia legislativa que se han producido en nuestro país, situándose con ellos a la vanguardia en este sentido.

2.4.1. Reglamento (UE) 2019/881 del Parlamento Europeo y del Consejo, de 17 de abril de 2019

Relativo a ENISA (Agencia de la Unión Europea para la Ciberseguridad) y a la certificación de la ciberseguridad de las tecnologías de la información y la comunicación y por el que se deroga el Reglamento (UE) nº 526/2013 («Reglamento sobre la Ciberseguridad»). En este Reglamento se disponen las nuevas competencias, tareas y estructura que tendrá la agencia ENISA, creada en 2004 y con sede en Grecia, que tiene como objetivo principal contribuir a la protección de los datos que circulan en Europa.

Con esto se trata de establecer un marco para la certificación en materia de ciberseguridad para que se puedan realizar evaluaciones a nivel comunitario de los conocimientos en este respecto, dado que actualmente existe una escasez de profesionales especializados en este ámbito. Entre las más importantes están: *Certified Ethical Hacker*, *CISM – ISACA*, *CISA – ISACA*, etc. Cabe destacar que España es uno de los países más avanzados en cuanto a certificaciones de ciberseguridad, con un organismo propio denominado Agencia española de Certificaciones de Ciberseguridad (ACC), creada en 2012 y de las mejores valoradas en la Unión Europea.

2.4.2. Directiva (UE) 2016/1148 del Parlamento Europeo y del Consejo, de 6 de julio de 2016

Traspuesta al ordenamiento jurídico español a través del Real Decreto-ley 12/2018, de 7 de septiembre. A través de este decreto se busca establecer la coordinación entre los organismos encargados de regular la seguridad de las redes y de los sistemas de información.

Esta también permitirá crear la Estrategia de Ciberseguridad Nacional, que estará formado por los CSIRT de Referencia. En consecuencia, se define al CCN-CERT como coordinador nacional. De esta manera, se establece una red institucional cuya finalidad será mejorar la seguridad frente a posibles ciberamenazas (Real Decreto-ley 12/2018, 2018).

En una entrevista personal, Eva Cañete (Unicaja Banco) expresa que "la normativa NIS por primera vez nos beneficia porque se exige la notificación de los incidentes de ciberseguridad, y eso nos puede ayudar a tener una base de datos, a tipificar incidentes y al final contribuye a tomar la temperatura del nivel de riesgo que tienen las empresas" (Google y The Cocktail Analysis, 2019).

2.4.3. Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales (LOPD-GDD)

Esta Ley aprobada en el Congreso de los Diputados trata de adaptar el ordenamiento jurídico interno de España al Reglamento General de Protección de Datos (RGPD), el Reglamento europeo aprobado en mayo de 2018, cuyo objetivo es alcanzar el mayor grado posible de protección de datos personales, en tanto que con él se protegen los datos de los ciudadanos europeos en todas las organizaciones, con independencia de que estas procedan o no de países miembros de la Unión Europea.

Como novedades cabe resaltar que, a partir de la inclusión de esta Ley, todos los consumidores debían ser debidamente informados de manera clara y entendible sobre el uso que se hará de sus datos (Ley 3/2018, 2018: art. 6). Además, se estableció la expresa prohibición de enviar información comercial sin previo consentimiento. Los principales principios de esta Ley son los siguientes (Agencia Española de Protección de Datos, 2019):

i) Principio de "licitud, transparencia y lealtad", mediante el cual los datos deben ser tratados de esta manera.

ii) Principio de "limitación de la finalidad", que tiene como resultado que los datos deban ser manejados con una o varias finalidades que hayan sido previamente determinadas, con lo que se prohíbe que estos sean usados para fines no legítimos y explícitos al usarlos para otro fin.

iii) Principio de "minimización de datos", es decir, que los datos sean limitados y adecuados para el fin que se persigue con ellos.

iv) Principio de “exactitud de datos”. En este principio se expresa que los datos deben ser exactos, es decir, que estén actualizados y que sean completos.

v) Principio de “limitación del plazo de conservación” de los datos, ya que estos pueden usarse de manera ilimitada en el tiempo. Así, una vez que se haya conseguido la finalidad propuesta los datos deberán ser eliminados de manera que, al menos, no sea posible que se identifique a los usuarios.

vi) Principio de “integridad y confidencialidad”, por el que los datos deberán ser tratados de manera segura para garantizar la privacidad de los usuarios.

Además, coincidiendo con la opinión de García (2019), las obligaciones de mayor importancia de la norma son las que se mencionan a continuación:

- Obligación, por parte del responsable del tratamiento, de notificar a los reguladores cualquier brecha de seguridad que se produzca en cuanto a los datos personales en un plazo máximo de 72 horas.
- En relación con el anterior, se obliga al responsable a comunicar dicha violación de seguridad a las personas afectadas, siempre que suponga para ellas un riesgo en cuando a derechos y libertades (Ley 3/2018, 2018: art. 73.s.).

Por otra parte, es importante saber que el silencio de los usuarios en cuanto a la otorgamiento del control de sus datos no implica consentimiento, por lo que debe haber un consentimiento explícito (Ley 3/2018, 2018: art. 6.2.). Cualquiera de las infracciones que se cometan en este respecto estarán sujetas a multas y sanciones por incumplimiento de la normativa, pudiendo alcanzar un total de $20,000,000 o el 4% del negocio total al año.

“En mi opinión esta ley ha entrado en vigor tarde, hubiera sido necesario tener una Ley específica en materia de ciberriesgos mucho antes, pero es cierto que las medidas que en ella se recogen han sido acertadas y adecuadas si con ello se consigue elevar el nivel de concienciación en ciberseguridad de las empresas que operen no solo en España, sino también en la Unión Europea. Es muy importante que con el desarrollo de las nuevas tecnologías y la continua innovación que se da en los mercados, los usuarios nos encontremos lo más protegidos posibles respecto a la seguridad de nuestros datos personales y nuestra privacidad” García (2019).

Sin duda, esta normativa ha supuesto un antes y un después en cuanto a concienciación, toda vez que una mayoría de personas hasta entonces no fueron del todo conscientes de la finalidad a la que se iban a destinar sus datos. Esto cambió cuando en las páginas webs se empezó a solicitar un permiso expreso informando de la misma, pudiendo el usuario aceptar o rechazar según su criterio y teniendo, por ende, el poder de decidir en mayor medida sobre su privacidad.

2.4.4. Estrategia Nacional de Ciberseguridad de 2019

Haciendo un análisis de la situación actual:

"Estamos hiperconectados. Nuestros negocios, estudios, salud y relaciones sociales. Nuestra vida. Y, sin embargo, no tenemos sensación de riesgo. Dudamos de si lo que ocurre en la Red es realidad o ficción. Dejamos nuestra existencia al descubierto" (Rodríguez, 2021).

Tanto es así que, como se ha expuesto anteriormente, el *ransomware* WannaCry dejó en evidencia la necesidad de contar con una estrategia común para dotar al Gobierno de una mayor defensa y protección en caso de ciberguerras, aspecto que ha resultado de gran utilidad en la crisis provocada por la pandemia del COVID-19 pues, aprovechando la inestabilidad y la debilidad, se han registrado numerosos ataques de hackers contra hospitales y farmacéuticas con el intento de golpear al sector sanitario (Rodríguez, 2021). Así, la Estrategia Nacional de Ciberseguridad recoge la necesidad de mejorar la seguridad en España en este sentido, además de detallar las medidas destinadas a este propósito. Sus principales líneas de acción son (Orden PCI/487/2019, 2019: cap.4):

i) Reforzar las capacidades ante amenazas cibernéticas.
ii) Garantizar la seguridad y resiliencia de los activos estratégicos.
iii) Investigación y persecución de la cibercriminalidad.
iv) Impulsar la ciberseguridad de ciudadanos y empresas.
v) Potenciar la industria española de ciberseguridad.
vi) Seguridad nacional.
vii) Desarrollar una cultura de ciberseguridad.

En conclusión, todo este esquema institucional y legislativo permite que España se sitúe a la vanguardia tecnológica en cuanto a ciberseguridad se refiere, lo cual nos permite evitar y mitigar los posibles y numerosos impactos adversos que tendrían los ataques cibernéticos en un ambiente ya de por sí complicado y debilitado por la situación que se da hoy día.

De esta forma, el único punto pendiente a reforzar sería en referencia al último de los puntos de la línea de acción, es decir, el conseguir una mayor campaña de concienciación para los usuarios por parte de las autoridades competentes, pues a pesar de saber los peligros que entrañan las redes, como se ha visto la mayoría de ellos no son del todo conscientes de que este tipo de problemas existe y que en cualquier momento pueden ser víctimas de los mismos. Del mismo modo, las iniciativas encaminadas a una mayor protección tampoco cuentan con la repercusión que debería tener este tipo de avances.

Igualmente, una vez detectado el riesgo mediante las diferentes estrategias que se han ido comentando es esencial contar con un proceso que permita poner el foco sobre la gestión del mismo.

2.5. EL PROCESO DE GESTIÓN DE RIESGO

Hasta ahora, el presente trabajo se ha centrado en poner en contexto una situación que, a pesar de no ser del todo notoria, coexiste con el día a día de todos dado el alto grado de dependencia que se da, tanto en individuales como en organizaciones, con las tecnologías de la información. En consecuencia, con el fin de concienciar e informar al lector, también ha sido definido el concepto de ciberataque y sus principales extensiones, citando el ejemplo más reciente que es el WannaCry.

Para proseguir con el análisis teórico, a continuación se expondrá el procedimiento tradicional de gestión de riesgos usado en las entidades a través de una normativa internacional llamada ISO 31000. Esta normativa puede ser aplicada a cualquier tipo de empresa y define la gestión de riesgos como “todas aquellas acciones coordinadas para dirigir y controlar los riesgos a los que puedan estar abocadas las organizaciones” (ISO Tools Excellence, 2017).

Además, los pasos a seguir para que esta gestión de riesgos sea efectiva serán (ISO Tools Excellence, 2017):

- Definición de objetivos. Con esta primera etapa, se deja claro qué se quiere conseguir a través de la gestión de riesgos; es imprescindible la herramienta del presupuesto de la compañía con el fin de asignar los recursos correspondientes a tal fin.

- Nombramiento de responsables. Todo proyecto debe tener "cabezas" visibles que sean las encargadas de asumir la responsabilidad del mismo. Para ello, las entidades pueden externalizar el proceso de selección a una firma independiente.

- Identificación de los riesgos. Se puede relacionar el riesgo con la incertidumbre que pueda provocar la posibilidad de que la privacidad de sus datos sea vulnerada frente a ciberataques con su consecuente impacto negativo en la entidad.

- Análisis de riesgos, con lo que se persigue hacer una clasificación de riesgos en función de su naturaleza, siendo necesario definir criterios cuantitativos o cualitativos.

- Definición de las respuestas a los riesgos. La normativa explica cinco estrategias para gestionar riesgos:

    i) Supresión del riesgo, consistente en eliminar el riesgo completamente. Esto se puede lograr a través de la planificación y la prevención.

    ii) Transferencia del riesgo entre diferentes empresas que pertenezcan a un mismo conjunto. Esto se logra al redirigir el riesgo al aceptar la otra parte el asumirlo.

    iii) Mitigación del riesgo. La clave de este proceso no es la eliminación del riesgo, sino reducirlo. Así, en caso de acaecer el suceso contingente la entidad tendrá pérdidas menores a las que hubieran ocurrido en caso de no adoptar esta posición.

iv) Explotación del riesgo. Hay algunos riesgos que son susceptibles de ser aprovechados por las organizaciones con el fin de sacarles el máximo rendimiento posible. Un ejemplo claro de esta circunstancia sería la crisis provocada por la pandemia, con la que han surgido oportunidades a través de la amenaza que supone.

v) Aceptación del riesgo. Son riesgos que por su naturaleza podrían convivir con la empresa al no suponer un problema de cara a cumplir con los objetivos definidos en un principio. Se muestra el ejemplo de una empresa que opere en Venecia, que lleva emparejado un riesgo de inundación que podría ser asumible a través de protocolos que agilicen la reacción en caso de suceder.

3. EL CIBERSEGURO, HIPÓTESIS DE PARTIDA

La gestión del riesgo cibernético en una empresa no es una tarea que a priori se pueda definir como sencilla, pues existe en una "revolución tecnológica" en la que predomina la hiperconectividad, la cual es definida como un concepto que integra la condición humana actual, en tanto que estamos permanentemente en contacto con la información a través de diferentes dispositivos como radio, televisión, Internet y teléfonos móviles (Gabriela Paoli, 2020). Así, de igual manera que ha aumentado este contacto con la información también lo ha hecho la exposición de las organizaciones a los riesgos que conlleva, que como se ha visto deben ser correctamente gestionados.

Aunque cada vez existe una mayor concienciación y prevención en lo que se refiere a los ciberriesgos, las empresas no deberían pensar que son capaces de llevar a cabo esta labor completamente por sí solas, pues hay contingencias que no podrán ser controladas y que terminarán teniendo un impacto que deberá ser cubierto para que este tenga la menor repercusión posible. En base a ello, en los últimos años ha sido también notable el auge de los ciberseguros como herramienta de protección contra eventos no esperados que derivan en grandes impactos para las empresas.

A diferencia de lo que ocurre con otros tipos de seguro, en el caso del ciberseguro no existe un procedimiento común, pues cada aseguradora ofrecerá unas coberturas diferentes, siendo por ello por lo que será complicado comparar varias ofertas y habrá pólizas "a la carta". En varias entrevistas para el diario Expansión, De las Casas (2019) recoge el testimonio de varios especialistas como Salvador Molina, presidente de Foro Ecofin, que afirma que "la ciberseguridad no puede ser solo una cuestión de indemnizaciones para el sector de los seguros, sino un reto de prevención, con un asesoramiento previo a las coberturas que analice y cierre vectores de riesgo en las empresas clientes" (De las Casas, 2019) y Alan Abreu, responsable de riesgos cibernéticos de Hiscox, que concluye que "las probabilidades de sufrir un ciberataque hoy son mucho mayores que un robo o un incendio, y eso es algo que las empresas deben tener en cuenta" (De las Casas, 2019).

Sin ir más lejos, se puede encontrar un ejemplo de esta última afirmación en la actualidad reciente, considerando que el Servicio Público de Empleo Estatal (SEPE) recibió en marzo de este mismo año un ataque cibernético a través de un fichero *spam*, el cual encriptó la información de los ordenadores y bloqueó la actividad durante días. Como resultado, la consecuencia directa podría haber tenido grandes implicaciones, y más aún en una situación tan delicada como el pago de los ERTE; habida cuenta de la ya de por sí compleja situación derivada de la pandemia. Así, el principal sindicato de la administración pública, la Central Sindical Independiente y de Funcionarios (CSIF), recordó que debido a esta incidencia se retrasará aún más la gestión de las citas para tramitar los expedientes de regulación de empleo, afectando esta circunstancia a más de 900.000 personas en nuestro país (Magallón, 2021).

3.1. INTRODUCCIÓN AL SEGURO PARA CIBERRIESGOS

Con el fin de introducir las nociones básicas sobre el contrato de ciberseguro se puede recurrir a la definición original de un seguro estándar, que sería el "antídoto" a un ciberriesgo o a una necesidad de cualquier tipo que se produzca a consecuencia del mismo y, por tanto, pueda resultar desfavorable. Así, este resarcirá al tomador de la póliza frente a un evento incierto, en este caso un ataque cibernético.

A cambio de este servicio se pagará una prima, la cual será modelizada en el siguiente capítulo, y se organizará un contrato de seguro en el que ambas partes quedarán vinculadas en virtud de unas condiciones determinadas que variarán según las coberturas que se incluyan. Guiándonos en base al reporte de Thiber (2016), se diferencian dos grandes tipos de coberturas, a saber:

3.1.1. Coberturas básicas

3.1.1.1. *Third-party coverage*

El servicio de cobertura frente a terceros no cubriría a la entidad directamente sino a un tercero, como puede ser un cliente que resultase perjudicado por un error de la compañía. Esta situación puede darse al filtrarse datos de carácter personal que puedan resultar en un ciberataque para ellos. Otros posibles escenarios donde podría actuar este ciberseguro serían (Thiber, 2016):

- Al traspasar, de forma no intencionada, un virus a un cliente a través del correo electrónico.
- Si se recomendara a un cliente un servicio que no fuera seguro.
- Uso de contraseñas sencillas de adivinar de cara a ciberdelincuentes que les permitieran entrar en las bases de datos de la compañía.

3.1.1.2. Procedimientos regulatorios

En este tipo de procedimiento se intenta dar cobertura frente a acciones legales emprendidas por el organismo correspondiente por el incumplimiento de la Ley de protección de datos que se veía anteriormente, lo cual traería como consecuencia una serie de sanciones que requieren de asesoramiento legal.

3.1.1.3. Gastos de gestión de incidentes

Sirven para mitigar los gastos referentes a la externalización de servicios como, por ejemplo (Thiber, 2016):

- Gastos forenses para investigar el origen del ciberataque y el abarque de los sectores afectados.
- Gastos de asesoramiento legal para hacer frente a las consecuencias judiciales derivados de los mismos.
- Gastos de comunicación y/o gestión del riesgo reputacional.
- Gastos de servicios prestados a los afectados, como contratación de servicios de atención de llamadas.

3.1.2. Coberturas opcionales o complementarias

3.1.2.1. *First-party coverage*

La cobertura de daños propios, o *first-party coverage*, sería la que correspondería a la empresa en el momento de ejecutar su póliza de seguro para pedir la correspondiente indemnización por acaecer uno de los eventos cubiertos, siendo la tarificación de este tipo de contrato será el objeto principal de estudio del presente trabajo. De esta forma, el objetivo del ciberseguro sería disminuir el impacto financiero por las pérdidas por ciberataques, e incluiría servicios de protección contra eventos como (Insureon, s.f.):

- Destrucción de datos intencionadamente.
- Ataques DDoS.
- Infección de virus de todas las categorías en el drive de la compañía.
- Daños contra el *hardware*.
- Destrucción accidental de bases de datos.
- Caída del servidor por sobrecarga eléctrica.

3.1.2.2. Responsabilidad civil de medios digitales

Esta sección ofrece cobertura frente al contenido que se publique en los sitios web de la organización. Entre los perjuicios que pueden dar lugar a este tipo de reclamaciones se encuentran "diversos motivos: desde la invasión de privacidad, calumnia y difamación a terceros hasta la vulneración de propiedad intelectual o marcas cuando se publican contenidos que pueden estar protegidos por derechos de propiedad intelectual de dichos terceros" (Thiber, 2016).

Además, cabe resaltar que no solo se puede recurrir a una póliza de ciberseguro para protegernos del riesgo frente a los delitos en la red, sino que también se puede hacer cobertura de cualquier evento que ocasione un impacto financiero negativo, incluyendo el riesgo reputacional, siendo necesaria su previa negociación de cara a incluirlo en la póliza de seguro. A la vez, existen riesgos que no se pueden asegurar en estas pólizas, como riesgos naturales o riesgos como los de incendio o explosión, así como actos deshonestos por parte del asegurado e infracción de secretos comerciales, entre otros.

3.1.3. Perfil de los potenciales clientes

Si bien toda organización es susceptible de ser víctima de un ataque cibernético, hay algunas que por el volumen de datos que manejan diariamente son más vulnerables que otras. Entre las principales afectadas estarían las instituciones financieras y aseguradoras, las teleco, y el sector sanitario y energético (Thiber, 2016), tal como se aprecia en la *Tabla 3.1*:

Sector	% Coste per cápita de los ciberincidentes
Sanitario	21%
Servicios financieros	17%
Retail	13%
Otros	11%
Tecnología	9%
Servicios profesionales	8%
Restauración	4%
ONG	4%
Hostelería	4%
Energía	2%
Medios de comunicación	2%
Manufactura	1%
Juego online y casinos	1%
Entretenimiento	1%
Transporte	1%
Telecomunicaciones	1%

Tabla 3.1 Distribución de siniestros y reclamaciones por sector (2016)

Fuente: Elaboración propia con datos de Thiber (2016)

A continuación, se comentará la situación de varios de estos sectores.

3.1.3.1. Instituciones financieras y aseguradoras

Este sector, sin duda, representa uno de los más importantes para el sostenimiento de las economías a nivel mundial, ya que es el que puede manejar mayor cantidad y volumen de datos de todo tipo, desde nombres y apellidos hasta tarjetas de crédito y direcciones. Esto, unido a un gran crecimiento de las nuevas aplicaciones de la mano de las *fintech*, hace que sea de vital importancia el que estén protegidas contra este tipo de ataques.

3.1.3.2. Empresas de telecomunicaciones

Estas empresas, como sucedía con Telefónica en el caso WannaCry, son especialmente sensibles en caso de recibir un ciberataque también por el tipo de datos que controlan, puesto que cada vez está más extendido el pago a través de los dispositivos electrónicos. De igual forma, este tipo de compañías están obligadas por normativa a reportar cualquier tipo de incidente en lo referente a violaciones de seguridad de datos y, en caso de no hacerlo, se expondrían a multas elevadas que pueden terminar en litigios si un tercero termina perjudicado, hecho por el que un ciberseguro es una opción recomendable para ellas (Esteve, 2015).

3.1.3.3. Sector sanitario

El verse afectados hospitales y, en general, toda administración sanitaria por un ataque de esta índole supondría que datos de toda clase se vieran comprometidos, como nuestros hábitos de vida, diagnósticos, enfermedades, etc. Así, además de cumplir con la Ley de protección de datos es menester que este sector tome medidas de protección adicionales, más aún si cabe en el momento en el que vivimos actualmente. Igualmente, los centros sanitarios han sido el objetivo principal de los ciberdelicuentes desde que empezó la pandemia y "muchas están sobreestimando su capacidad para proteger unos entorno que, por el momento, no conocen bien" (Alonso, 2021).

3.1.3.4. Sector energético

La industria energética ha ido adquiriendo herramientas tecnológicas cada vez más avanzadas, pero esto significa que se ha vuelto cada vez más frágil en términos de ciberseguridad. La aparición de determinados ataques en este sector podría incluso poner en jaque la seguridad de los ciudadanos, por lo que deben ser conscientes de que se hallan en el punto de mira. De hecho, en 2019 se identificaron nuevos y preocupantes ataques DDoS en grandes potencias como EEUU. Así, estos nuevos peligros han propiciado que los reguladores promuevan medidas destinadas a proteger aún más este sector.

3.1.4. Intervinientes en el contrato de la póliza de seguro

Para entrar en este mercado hay que ser empresario o asegurador, y a este último se le exigirá que la actividad típica sea el intercambio de una prestación presente y cierta, que es la prima de seguro, por una prestación futura e incierta, que será la prestación. Así, el asegurador se comprometerá a la satisfacción de la indemnización, la suma asegurada, por lo que es un compromiso de pago en el caso de ocurrencia del riesgo asegurado.

El asegurador debe estar en perfectas condiciones para poder asumir la cobertura cuando se produzca el riesgo cubierto, otorgando al asegurado la seguridad de que responderá por él para hacer frente al mismo, dado que le resarcirá cuando se vea afectado. Su actividad es la gestión de los riesgos de sus asegurados, de manera que el asegurador será en todo momento perfecto conocedor de los mismos y su negocio irá enfocado a una gestión prudencial y adecuada de los riesgos transferidos.

Por otro lado se encuentra el tomador, que espera que aquel evento sea garantizado cuando se produzca el riesgo y espera y desea que el asegurador cumpla con esa promesa de pago futuro cuando acaezca el evento objeto de cobertura.

Además, uno de los actores más importantes serán las reaseguradoras, que asumirán ciertos riesgos que sean transferidos por las aseguradoras y que será de vital importancia para el tratamiento de los siniestros.

3.2. EL REASEGURO EN EL SECTOR

A través del contrato de reaseguro se unirán los dos actores principales en esta modalidad: el cedente (la aseguradora) y la reaseguradora. Así, la cedente pagará una prima al reasegurador para que a cambio cubra una parte de las pérdidas en las que incurrirá en el futuro. Hay tres momentos en los que destaca la importancia de la relación entre cedente-reaseguradora, como son (García, 2019):

3.2.1. Formalización del contrato de reaseguro para el ramo de ciberriesgos entre la reaseguradora y la cedente

El papel del reasegurador es esencial a la hora de lanzar un nuevo producto al mercado, una vez que este puede asesorar a la cedente sobre qué tipo de contrato es el más favorable en base a sus necesidades, teniendo en cuenta el equilibrio existente entre la retención de la cedente y la cesión al reaseguro. Para este asesoramiento la reaseguradora se basa en el juicio experto derivado de la experiencia. En este ensayo se estudian las características principales de los dos ramos principales del reaseguro que más se dan en la actualidad (Instituto de Ciencias del Seguro, 2010): el contrato proporcional y el no proporcional.

3.2.1.1. Contrato proporcional

En él, el riesgo que traspasa la aseguradora a la reaseguradora es proporcional al riesgo que asume la primera, quedándose por tanto con la proporción restante de los que no han sido cedidos, de ahí el nombre de este tipo de contratos. De esta forma, la reaseguradora se hace cargo de una porción de los riesgos que ha suscrito la aseguradora con sus clientes y se hará cargo tanto de los siniestros que estos tengan como de las primas que pagan. Existen dos tipos de contratos proporcionales:

- Cuota parte, o *quota-shared*: es un tipo de contrato proporcional se representa como "X" la pérdida posible en un contrato de ciberseguro. Del contrato cuyo riesgo cubre la compañía que vende la póliza, le va a ceder una fracción fija de riesgo al reasegurador, "α". Así, el reasegurador pagará "$X\alpha$" (3.1) mientras que la aseguradora pagará la parte restante, más conocida como parte retenida, "$(1-\alpha)X$" (3.2). Esto se puede encontrarlo de forma práctica con el coaseguro.
- Contrato de excedente: de metodología similar al anterior, pero aplicando el factor "α" únicamente para los riesgos que superen un límite fijado previamente. De esta forma, los riesgos que pertenecen a la aseguradora después de este reparto se denominarán "plenos de retención", mientras que los pertenecientes a la reaseguradora se llamarán "excedentes"; a mayor riesgo, menores plenos de retención tendrá la aseguradora (Instituto de Ciencias del Seguro, 2010).

A modo de resumen se diferencian las principales ventajas y desventajas de estas modalidades de contrato proporcional en la *Tabla 3.2*:

Ventajas	Desventajas
Permiten establecer un equilibrio en la cartera de riesgos que retiene por cuenta propia	Dificultad de medir el nivel de retenciones
Fácil de calcular a nivel estadístico	No protegen eficazmente al reasegurado contra un cúmulo de daños por siniestros generalizados como catástrofes naturales o algo que afecte a la vez a muchos asegurados
No hay efecto de la inflación sobre "**α**", pues si se ha reasegurado este porcentaje de cada siniestro, no importa en qué momento del tiempo sea	Coste elevado, ya que a veces se usa el reaseguro para ceder riesgos que en realidad no serían necesarios, dependiendo de la cantidad. Solo interesaría para cuantías elevadas

Tabla 3.2. Ventajas y desventajas del contrato de reaseguro proporcional

Fuente: Elaboración propia con datos del Instituto de Ciencias del Seguro (2010)

3.2.1.2. Contrato no proporcional

En esta modalidad de reaseguro se sigue la misma idea que en el proporcional, es decir, se aplicará una división a cada uno en cuanto a los siniestros acaecidos pero, en este caso, esta división no será en base a un porcentaje, sino que la reaseguradora pagará el exceso del siniestro que resulte por encima de cierto umbral previamente fijado.

De esta manera, si el siniestro queda por debajo de este límite, no deberá cubrirlo, con lo que se resolvería uno de los problemas de la modalidad anterior. Hay dos tipos de contratos no proporcionales:

- Exceso de pérdidas por riesgo, o *excess of loss*. Llevado a cabo a través de la franquicia y el límite a la suma asegurada, donde desde el punto de vista de la aseguradora cuando el importe del siniestro, "X", sea menor que el umbral que se fije, "d", la aseguradora se hará cargo de la totalidad del siniestro. Por otra parte, si el importe de la pérdida es superior la aseguradora solo pagará hasta dicho límite, corriendo el resto a cargo de la reaseguradora. Así, el riesgo para la primera quedaría truncado en "d". Por su parte, el reasegurador no pagaría nada en caso de que el siniestro quedara por debajo del umbral y se haría cargo del exceso en caso de superarlo. Definiendo Y_r como la siniestralidad a cargo de la aseguradora, con el límite a la suma, e Y_c como la de la reaseguradora, con la franquicia:

$$Y_r = \begin{cases} X, si\ X \leq d \\ d, si\ X > d \end{cases} \tag{3.3}$$

$$Y_c = \begin{cases} 0, si\ X \leq d \\ X - d, si\ X > d \end{cases} \tag{3.4}$$

- Exceso de siniestralidad, o *stop loss*. En esta subdivisión se cede a la reaseguradora la parte que exceda un umbral, al igual que en la anterior, pero sobre el total de la cartera. La diferencia con el *excess of loss* es que allí se cedía el importe individual de cada póliza, no de la cartera en general, por lo que habría que esperar al final del ejercicio para ver cuánto se cede en total. Así, en esta modalidad la cuantía a cargo de la reaseguradora se puede representar como:

$$S_d = \begin{cases} 0, si\ S \leq d \\ S - d, si\ S > d \end{cases} \tag{3.5}$$

De nuevo, en la *Tabla 3.3* se repasarán las principales ventajas y desventajas de los contratos de reaseguro no proporcionales:

Ventajas	Desventajas
Ahorro en gastos de administración de las cedentes por su simplicidad administrativa	Problemas de financiación de las cedentes
El asegurador puede retener más primas suscritas	Establecimiento del umbral límite, “d”
Mayor flexibilidad para la suscripción de riesgos	Dificultades en cuando al cálculo del precio de la cobertura

Tabla 3.3. Ventajas y desventajas del contrato de reaseguro no proporcional

Fuente: Elaboración propia con datos del Instituto de Ciencias del Seguro (2010)

3.2.2. Ayuda en la suscripción a la aseguradora: análisis y apetito del riesgo

Para definir la ayuda en la suscripción a la aseguradora se puede afirmar lo siguiente:

“La reaseguradora pretende ayudar a la cedente en su política de suscripción cuando esta lanza un nuevo producto asegurador en el mercado y desconoce qué riesgos debe admitir y cuáles rechazar, así como la tarifa que debe aplicar por insuficiencia de datos actuariales” (García, 2019).

Por otro lado, en línea con este autor uno de los grandes problemas en cuanto a la suscripción de nuevas pólizas es la falta de datos de siniestralidad y del número de pólizas, lo cual complica el encontrar una metodología exacta para cuantificar la prima de ciberseguro. En el presente ensayo se aplicará un método aplicando simulación mediante cópulas, pues en los ciberseguros no se puede aplicar la Ley de grandes números que se suele usar en el sector ni se tienen en cuenta los elementos típicos del reaseguro. Además, este es un seguro relativamente nuevo, con lo que es complicado de modelizar al aparecer cada día nuevas amenazas que pueden hacer variar las coberturas. En consecuencia, este trabajo podría contribuir a ilustrar cómo usar los datos empíricos disponibles para llevar a cabo el *pricing* de forma diferente a lo que se ha venido haciendo hasta ahora con modelos ad-hoc.

3.2.3. Ayuda en la tramitación de siniestros a la aseguradora

Dada la complejidad en la tramitación y valoración de los siniestros por lo comentado anteriormente, "podríamos decir que las reaseguradoras ejercen de guía metodológica a las compañías de seguros (…), ya que estas últimas a veces no disponen de los recursos técnicos suficientes para gestionarlo correctamente" (García, 2019).

Para concluir este apartado, una vez vistas las diferentes modalidades del reaseguro y sus implicaciones de cara a mitigar el riesgo excesivo que puedan sufrir las aseguradoras por el alto riesgo existente en el ciberseguro se introducirá cómo intentar aproximar su prima mediante un método diferente, apoyándonos en el enfoque con cópulas.

3.3. INTRODUCCIÓN A LA VALORACIÓN DEL CIBERSEGURO MEDIANTE CÓPULAS

El ciberseguro es, sin duda, un método eficaz de disminuir los riesgos que se han venido estudiando. Sin embargo, como se comentaba no existe una método estándar para tarificar su prima, por lo que en el enfoque que se propondrá se usará la simulación y las funciones cópula, basadas en el *paper* de Hearth y Herath (2011), las cuales introducen dos variables fundamentales a tener en cuenta para determinar la prima de ciberseguro: el número de ordenadores afectados y el importe de las pérdidas a consecuencia de los ataques.

De esta forma, este tipo de seguro no es comparable al resto, ya que la tecnología influye considerablemente en esta casuística (Hearth y Herath, 2011). En este sentido, los riesgos relativos a Internet son únicos en términos de localización y visibilidad y "las políticas tradicionales no abordan de manera integral los riesgos adicionales a los que hacen frente las empresas como resultado de ser parte de la economía digital" (Gordon et al., 2003; en Hearth y Herath, 2011).

En consecuencia, como se veía al finalizar el apartado anterior, en otros productos de seguro se recurre a metodologías actuariales de tarificación que aquí no serían aplicables. La razón principal de este argumento se encuentra en lo novedoso que es el uso de Internet, con lo cual no se dispone de un histórico completo de frecuencia sobre el que sustentar los análisis, por lo que se indagará únicamente en la severidad, o cuantía, del ciberseguro.

Una forma de hacer frente a esta peculiaridad es a través de las cópulas definidas por Sklar (Sklar, 1959), que nos servirán como técnica para medir la estructura de dependencia entre las funciones de distribución empíricas de nuestras variables aleatorias. Así, estas "permiten construir una distribución de probabilidad conjunta que representa [dicha] dependencia" (Novales, 2017).

De esta manera, la cópula podrá definirse como una "función de distribución multivariante definida sobre el cubo unidad $[0,1]^n$ con marginales" (Cintas del Río, 2007). En el caso concreto de las cópulas bivariantes, objeto de estudio, se puede decir que cumple con las siguientes propiedades (Novales, 2017):

- $C : [0,1]\ x\ [0,1] \rightarrow [0,1]$
- $C(u_1, 1) = u_1;\ C(1, u_2) = u_2$
- $C(u_1, 0) = C(0, u_2) = 0$
- Para todo u_1, u_2, v_1, v_2 en [0,1] con $u_1 \leq v_1$, $u_2 \leq v_2$, se tiene: $C(v_1, v_2) - C(u_1 - v_2) \geq C(v_1, u_2) - C(u_1, u_2)$

Con ello:

"La primera condición permite utilizar una cópula sobre los valores tomados por funciones de distribución. La segunda condición es el requisito de tener distribuciones marginales (…). Las otras dos condiciones son propias de toda función de distribución bivariante" (Novales, 2017).

Además, cabe introducir el Teorema de Sklar, según el cual dada una función de distribución bivariante $F(x_1, x_2)$ con distribuciones marginales F_1, F_2, existe una cópula C: $[0,1]\ x\ [0,1] \rightarrow [0,1]$ tal que $\forall x_1, x_2$ en R^2 se tiene (Novales, 2017):

$$\forall (x_1, x_2) \in R^2 \rightarrow F(x_1, x_2) = C\big(F_1(x_1), F_2(x_2)\big) \qquad (3.6)$$

Existiendo, por ende:

"Una cópula que toma sobre R^2 los mismos valores numéricos que la función de distribución $F(x_1, x_2)$. Decimos que C es la cópula de la distribución F (…) Si las distribuciones marginales son continuas, dicha cópula es única. Recíprocamente, si C es una cópula y $F_1(x_1)$, $F_2(x_2)$ son funciones de distribución univariantes, entonces (3.6) define una función de distribución bivariante con distribuciones marginales $F_1(x_1)$, $F_2(x_2)$" (Novales, 2017).

En esto último será en lo que se basarán los cálculos propuestos, al construir una función de distribución conjunta a través de dos funciones de distribución marginales y la cópula. Sobre ella se simulará con Monte Carlo una muestra aleatoria bivariante, que nos ayudará a obtener la distribución de la suma asegurada. Las cópulas serán interesantes por dos razones fundamentales (Hearth y Herath, 2011): (i) consideran dependencias no lineales, al no tratar marginales normales, y (ii) sirven de punto de partida para construir distribuciones bivariantes, que nos servirán para aplicar simulación de Monte Carlo. Para ello, previamente se determinará qué cópula de las existentes es la que mejor se ajusta a los datos: la cópula de Clayton (Clayton, 1978) o la de Gumbel (Gumbel, 1960), más conocidas como cópulas arquimedianas.

Por un lado, la cópula de Clayton tiene dependencia asimétrica en las colas, mientras que la de Gumbel tiene dependencia positiva en la cola superior. De esta manera, las cópulas de supervivencia serán el contrario de dichas cópulas, es decir, la parte restante hasta llegar a la unidad. Se puede observar de forma más intuitiva la representación de la cópula de Clayton a través de la *Figura 3.1* y *Figura 3.2*, ayudándonos del paquete de R *copula* (Hofert y Mächler, 2011):

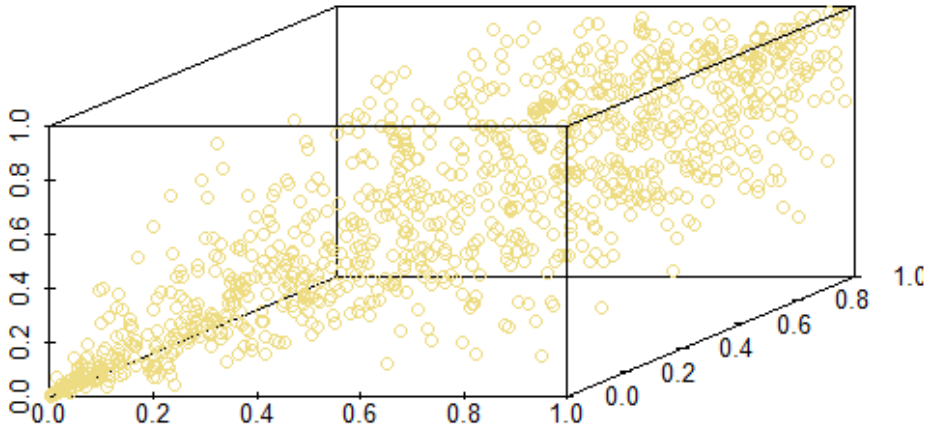

Figura 3.1 Cópula de Clayton

Fuente: Elaboración propia

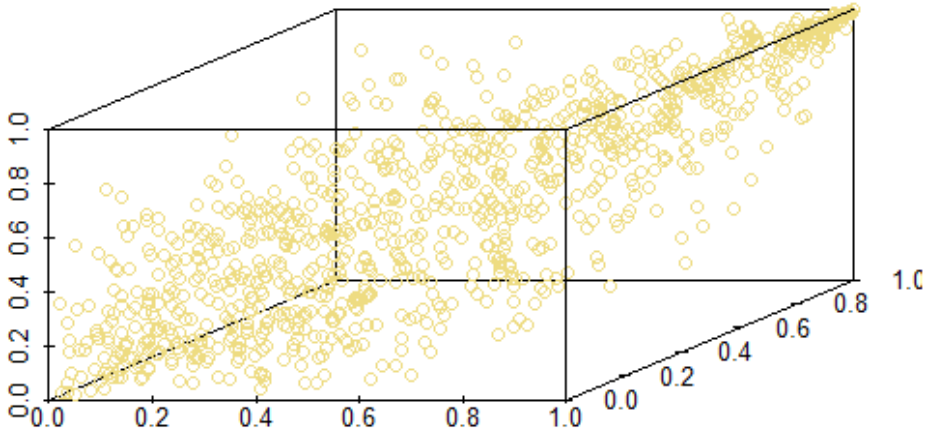

Figura 3.2 Cópula de supervivencia de Clayton

Fuente: Elaboración propia

Por otra parte, la *Figura 3.3* y la *Figura 3.4* nos ayudarán a interpretar la cópula de Gumbel:

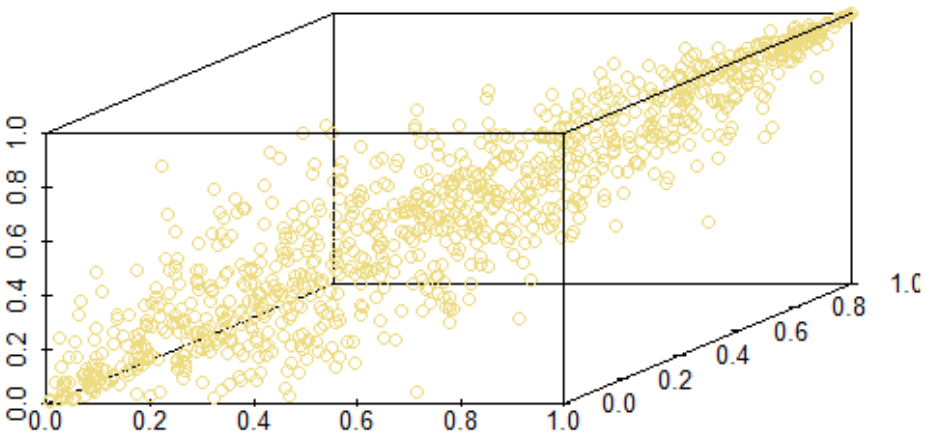

Figura 3.3 Cópula de Gumbel

Fuente: Elaboración propia

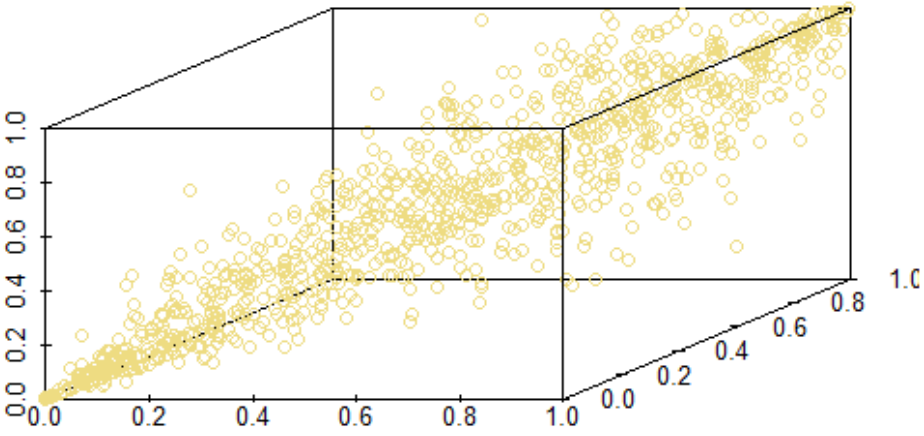

Figura 3.4 Cópula de supervivencia de Gumbel

Fuente: Elaboración propia

Como se verá a continuación, se buscará la cópula que mejor describa la estructura de dependencia entre las funciones de distribución empíricas de las variables aleatorias, que influirán en la determinación de la prima de ciberseguro.

4. MODELIZACIÓN Y PRICING DEL CIBERSEGURO, METODOLOGÍA Y RESULTADOS

4.1. MODELOS EPIDEMIOLÓGICOS

Antes de analizar las funciones cópula, se hará un breve repaso de los procesos epidémicos con el fin de ver cómo se podría modelizar la propagación de un virus por Internet, los cuáles servirán de herramienta para anticiparnos ante posibles escenarios y así completar el caso de estudio. Para este apartado se analizará el estudio realizado por Bichara et al. (2013), cuyos cálculos han sido adaptados por Petrelli (2019). Así, estos muestran que, en general, en los modelos de contagio hay varios individuos que se pueden encontrar en diversos estados, a saber:

i) Susceptible (S), ordenadores que podrían contraer un virus determinado.
ii) Expuestos (E), es decir, ordenadores que han sido infectados por el virus pero que no han desarrollado "síntomas", con lo que no podrían transmitirlo al grupo S.
iii) Infectados (I), como consecuencia de haber estado expuestos. Estos podrían transmitir el virus al grupo S.
iv) Recuperados (R), es decir, que han pasado el virus y que tienen un grado de protección por ello.

En este caso se estudiarán modelos que no tienen por qué seguir este orden reglamentariamente. Por ejemplo, en el modelo SIS los ordenadores susceptibles (S) pueden pasar al de infectados (I) al contraer el virus, pero una vez que se recuperan volverán a ser susceptibles de contagiarse (S). En contra, en el modelo SIR los individuos susceptibles (S) pasan al grupo de infectados (I) y finalmente entran en el grupo de recuperados (R) al dejar de estar infectados consiguiendo la inmunidad.

4.1.1. Modelo SIS

En este modelo, el virus seguirá el proceso mostrado en la *Figura 4.1* (Bichara et al., 2015; en Petrelli, 2019):

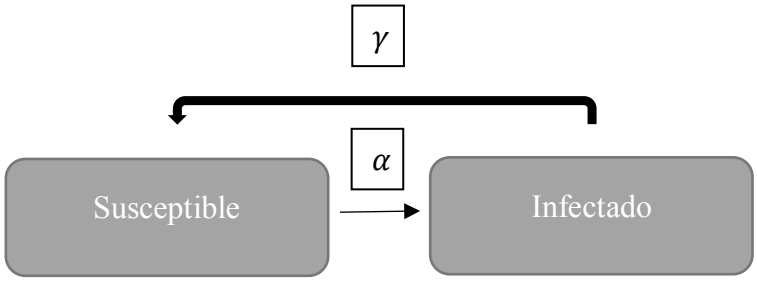

Figura 4.1 Esquema modelo SIS

Fuente: Elaboración propia basada en Petrelli (2019), con datos de Bichara et al. (2015)

Con lo que se destaca que los ordenadores podrán contagiarse a una tasa de infección de "α", mientras que volverán al estado "susceptible", es decir, se curarán, a una tasa de "γ". Siguiendo el principio de parsimonia, se aplicará la metodología más sencilla, con lo que se resumirán ambas variables en una sola: β, que será el porcentaje de ordenadores sanos que han sido afectados por el virus. De esta forma, el modelo resultante será el siguiente (Petrelli, 2019):

$$\frac{dS}{dt} = -\beta SI + \gamma I \tag{4.1}$$

$$\frac{dI}{dt} = \beta SI - \gamma I \tag{4.2}$$

Definiendo "N" como el total de ordenadores se puede decir, de forma trivial, que:

$$N = I + S \tag{4.3}$$

De lo cual, al despejar "I" y sustituir en la Ecuación (4.3) se obtiene que:

$$\beta S^2 - S(\gamma + \beta N) + \gamma N = 0 \tag{4.4}$$

Si nos fijamos en el procedimiento de este modelo, se entra en una especie de "bucle" en el cual los ordenadores susceptibles pasarán a estar infectados, volviendo de nuevo a ser susceptibles, y así indefinidamente. Mientras tanto, "N", el total de la población, definido en la Ecuación (4.3), permanecerá constante. Con ello, el virus no se extinguirá nunca, con lo que se deberá imponer la condición al modelo de que alguna de las derivadas de las Ecuaciones (4.1) y (4.2) sea 0 (Bichara et al., 2015).

"La Ecuación (4.1) es una ecuación cuadrática simple cuyas raíces son $S = \frac{\gamma}{\beta}$ y $S = N$, que son los estados de equilibrio para el número de individuos susceptibles. Si el estado de equilibrio de la población susceptible (S) es igual a "N", eventualmente todos serán susceptibles y nadie se infectará, por lo que la enfermedad se extinguirá; si por el contrario nadie se recupera de la infección (k = 0), finalmente todos estarán infectados y no habrá individuos susceptibles" (Petrelli, 2019).

Con esto, usando el *software* R se aplicará la simulación con las ecuaciones propuestas, suponiendo como hipótesis inicial que N = 100 (Petrelli, 2019). Para ello, nos ayudaremos del paquete *deSolve* (Soetaert, K. et al., 2009), que utiliza el método de Runge Kutta de paso variable para resolver las ecuaciones diferenciales de primer orden, y el paquete *simecol* (Rinke y Petzoldt, 2003) para simular los sistemas dinámicos en el ámbito ecológico. Así, considera unos valores iniciales de I = 2 y S = 98. Sin embargo, se observa que, sea cual sea el estado inicial, "S" llega siempre a $\frac{\gamma}{\beta}$, 20, o a "N", 100:

i) Simulación con $\gamma = 1$ y $\beta = 0.05$:

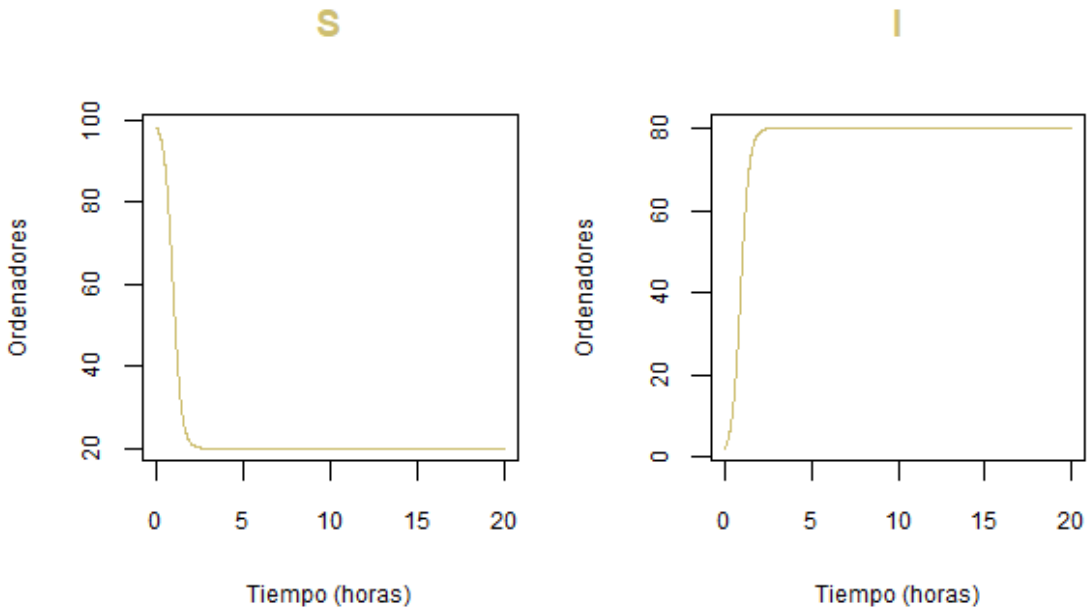

Figura 4.2 Simulación modelo SIS con parámetros $\gamma = 1$ y $\beta = 0.05$

Fuente: Elaboración propia basada en los cálculos de Petrelli (2019)

En este caso, en la *Figura 4.2* se puede observar que el número de ordenadores que se encuentran en estado susceptible cae de forma abrupta rápidamente, mientras que en el caso de los infectados sería el caso contrario, pues la tasa de infección es elevada.

ii) Simulación con $\gamma = 1$ y $\beta = 0.001$:

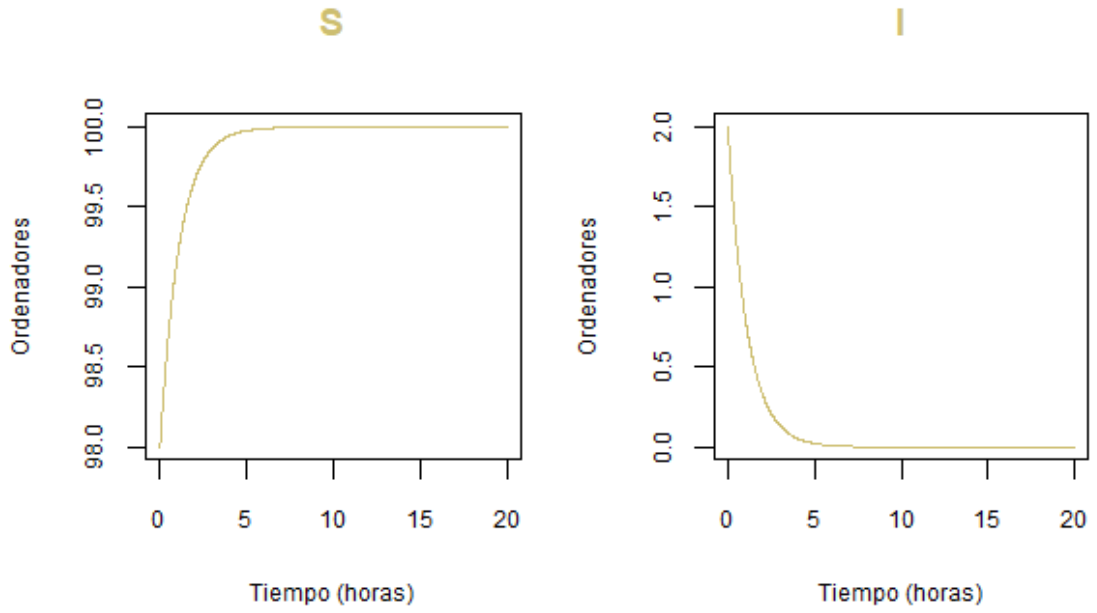

Figura 4.3 Simulación modelo SIS con parámetros $\gamma = 1$ y $\beta = 0.001$

Fuente: Elaboración propia basada en los cálculos de Petrelli (2019)

Por contra, en la *Figura 4.3* se comprueba que el efecto de disminuir el parámetro β repercutirá en un decremento de los ordenadores infectados, ya que el número de ordenadores en este estado cae a 0 de forma rápida, sucediendo lo contrario con los ordenadores en estado susceptible.

iii) Simulación con $\gamma = 1$ y $\beta = 0.01$:

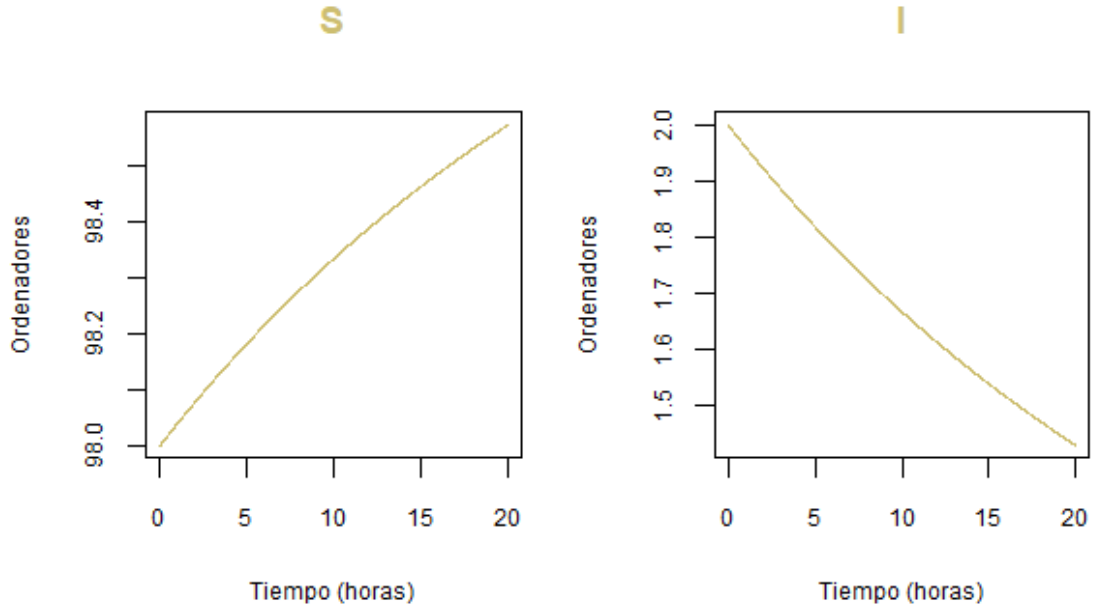

Figura 4.4 Simulación modelo SIS con parámetros $\gamma = 1$ y $\beta = 0.01$

Fuente: Elaboración propia basada en los cálculos de Petrelli (2019)

Al imponer un valor intermedio de β igual a 0.01, en la *Figura 4.4* se observa que el número de ordenadores en estado susceptible aumenta de forma casi lineal en el tiempo, lo contrario que sucede con los que están infectados.

Con todo, lo importante que se sustrae del análisis de este modelo es la demostración de los dos posibles estados en los que pueden estar los ordenadores, pues, como se ha observado, en cada unidad de tiempo una parte de los ordenadores infectados se "curarían" del virus y volverían al estado susceptible y viceversa.

4.1.2. Modelo SIR

Este modelo, propuesto por Li y Zou (2009), parte de las hipótesis de que los estados por los podrán transitar los ordenadores serán "susceptible", "infectado" y "recuperado", como se puede apreciar en la *Figura 4.5*:

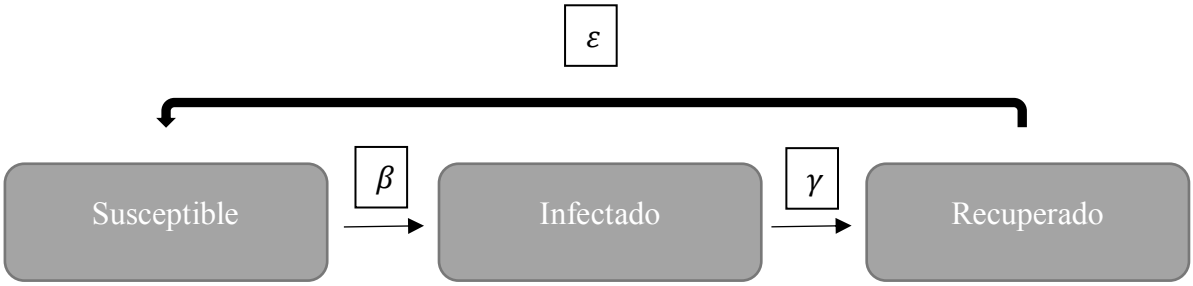

Figura 4.5 Esquema modelo SIR

Fuente: Elaboración propia basada en Petrelli (2019) con datos de Li y Zou (2009)

Se supondrá que los ordenadores se contagiarán a una tasa de infección de "β", se recuperarán a una tasa de "γ" y pasarán de nuevo al estado susceptible a una tasa "ε". De nuevo, el modelo resultante de ecuaciones diferenciales será el siguiente (Petrelli, 2019):

$$\frac{dS}{dt} = -\beta SI + \varepsilon R \tag{4.5}$$

$$\frac{dI}{dt} = \beta SI - \gamma I \tag{4.6}$$

$$\frac{dR}{dt} = \gamma I - \varepsilon R \tag{4.7}$$

A continuación, se analizán en la *Figura 4.6* las posibles tendencias que pueden seguir los ordenadores en los diferentes estados en el intervalo de 12 horas, con estados iniciales S = 98, I = 1 y R = 0:

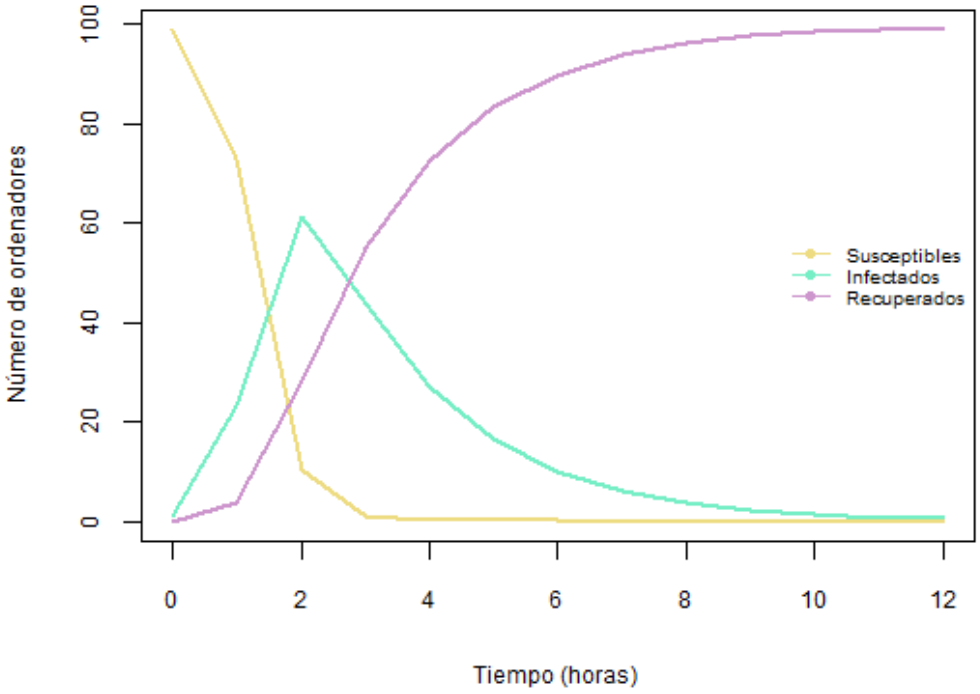

Figura 4.6 Simulación modelo SIR con parámetros $\gamma = 0.5$, $\beta = 0.04$ y $\varepsilon = 0$

Fuente: Elaboración propia basada en los cálculos de Petrelli (2019)

En primer lugar, por lógica y simplicidad, se ha impuesto la condición de que la tasa ε sea igual a cero, con lo cual un ordenador que se recupere adquirirá una "inmunidad" tal que no permita que vuelva al estado susceptible. Así, con los parámetros estudiados se observa que los ordenadores en este estado van cayendo rápidamente pasadas las dos primeras horas, quedando menos de 20 susceptibles y coincidiendo en tiempo con el pico de la curva de infectados, que alcanza un valor de 60 aproximadamente. Además, los ordenadores recuperados van aumentando progresivamente hasta estabilizarse en los 99 pasadas las 12 horas.

Una vez vista la modelización de la propagación de los virus a través de los modelos propuestos se pasará a la siguiente sección, donde se organizará la estructura que resulta esencial para realizar el *pricing* del ciberseguro, las funciones cópula.

4.2. FUNCIONES CÓPULA

Para acercar al lector sobre el funcionamiento de estas funciones se recurrirá a un ejemplo utilizando el paquete *MASS* (Yan, 2007) en R para generar varias muestras de una distribución normal multivariante de tres variables aleatorias. Se comenzará por la siguiente matriz inicial:

$$\begin{pmatrix} 1 & 0.2 & 0.1 \\ 0.2 & 1 & -0.4 \\ 0.1 & -0.4 & 1 \end{pmatrix} \tag{4.8}$$

Llegando a una distribución normal multivariante, de tres variables, mostrada en forma de matriz en la *Figura 4.7*, de la que se escoge para su representación las siete primeras filas de una de las simulaciones aleatorias:

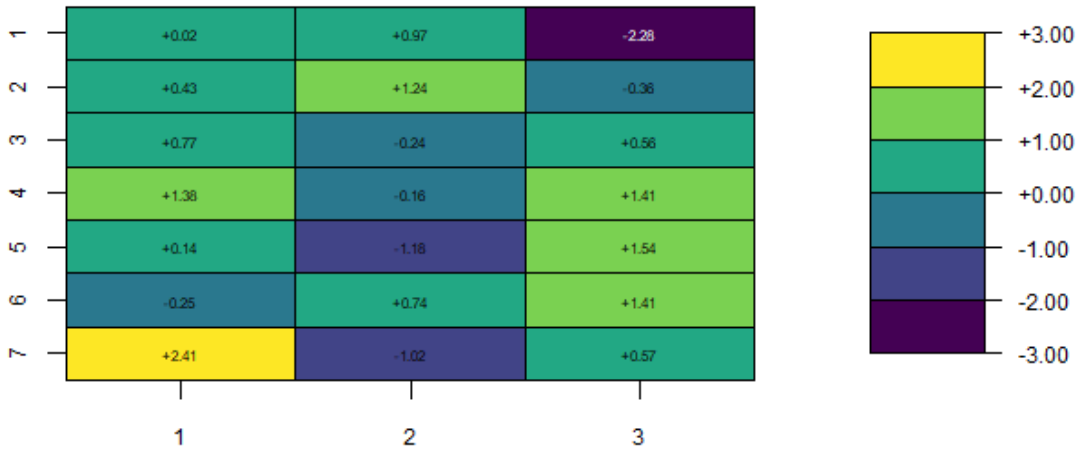

Figura 4.7 Simulación distribución normal multivariante

Fuente: Elaboración propia

El siguiente paso, una vez obtenida esta distribución hipotética, será ver la correlación existente entre las tres muestras simuladas usando el coeficiente de correlación de Spearman (Spearman, 1904), el cual permitirá comprobar la intensidad de la relación entre las mismas. Usaremos este coeficiente y no el de Pearson (Pearson, 1895) al no tratarse de variables con la misma varianza, que será una limitación que se verá más adelante. Así, se obtiene la siguiente matriz de correlaciones:

$$\begin{pmatrix} 1 & 0.18 & 0.10 \\ 0.18 & 1 & -0.38 \\ 0.10 & -0.38 & 1 \end{pmatrix} \tag{4.9}$$

De igual forma, en la *Figura 4.8* se representan las correlaciones a pares entre las tres muestras aleatorias obtenidas, siendo la diagonal la varianza de cada una de ellas:

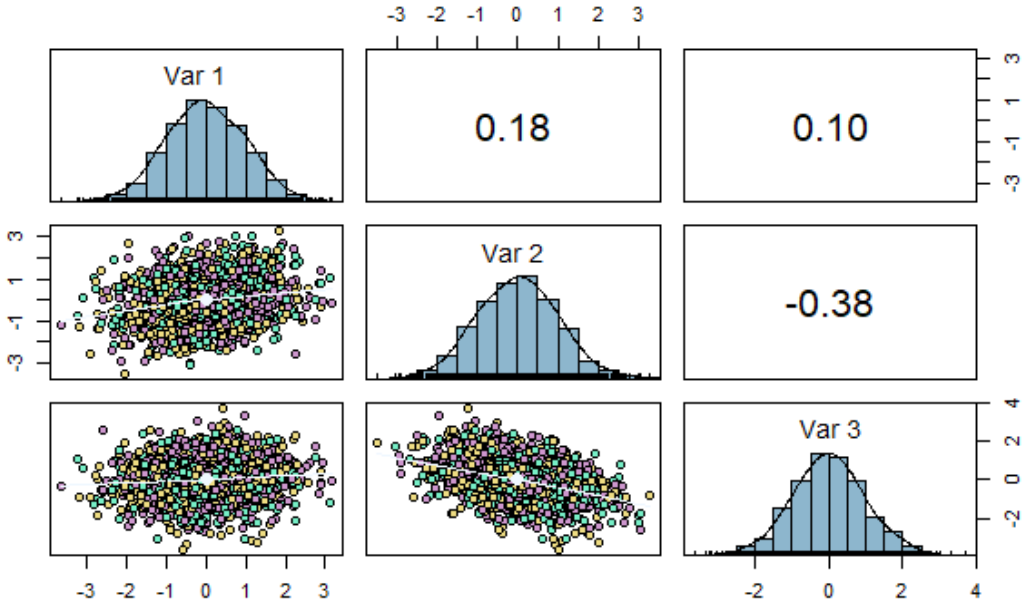

Figura 4.8 Correlaciones a pares entre las variables observadas

Fuente: Elaboración propia

Considerando esta muestra como la observada, el siguiente paso será buscar una manera de simular otra muestra, con simulación de Monte Carlo, siendo imprescindible para dar por válidos los resultados que posean una correlación similar. Para ello, en primer lugar se ajustarán diferentes distribuciones a las variables y, en segundo lugar, se usarán las cópulas para encontrar la estructura de dependencia de las mismas. Una vez se tenga esto, se obtendrá como resultado una distribución multivariante simulada con la que se aproximará la prima de ciberseguro.

Para poder aplicar las funciones cópula se transformarán las variables aleatorias a la distribución uniforme en el intervalo [0,1], que resultaría como sigue en la *Figura 4.9*. Además, se puede apreciar de forma más visual su representación gráfica en la *Figura 4.10*.

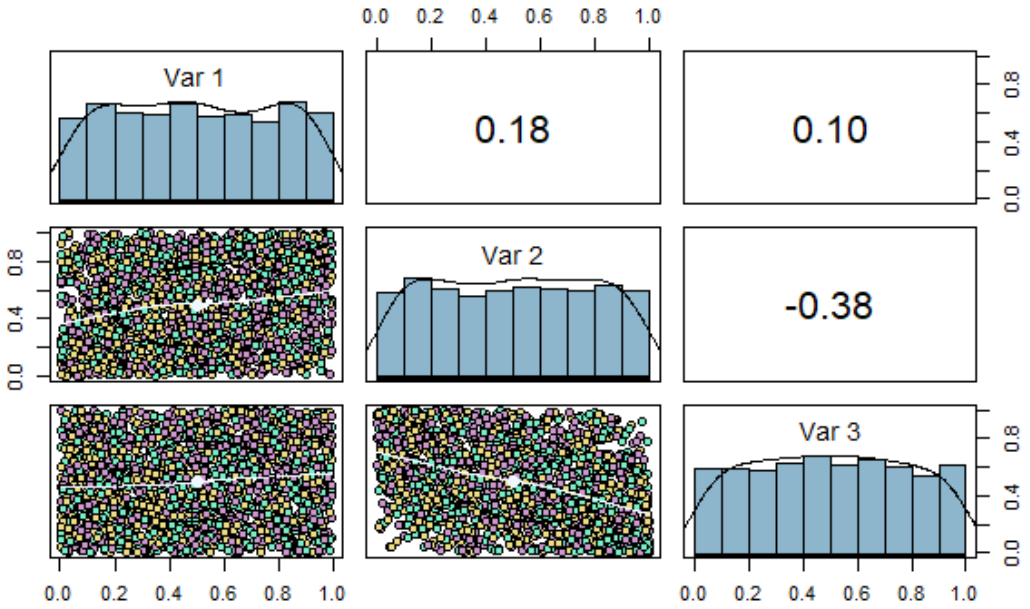

Figura 4.9 Correlaciones a pares entre las variables observadas transformadas a la distribución uniforme [0,1]

Fuente: Elaboración propia

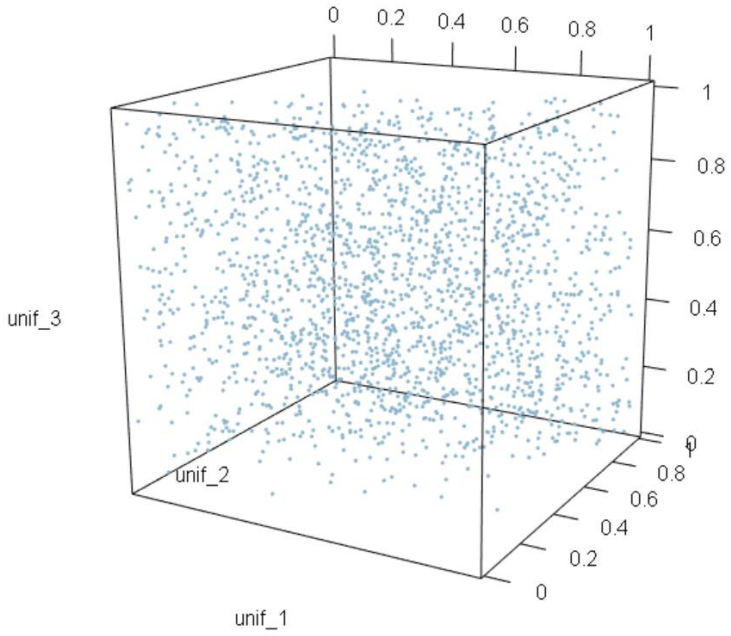

Figura 4.10 Variables simuladas transformadas a la distribución uniforme [0,1]

Fuente: Elaboración propia

Sin duda, el aspecto más destacable que se aprecia en la *Figura 4.9* es que, al cambiar la distribución de las variables, no se ha alterado la estructura de dependencia, dado que se puede observar que las correlaciones entre ellas siguen siendo las mismas.

A continuación, se escogerán las distribuciones marginales que se ajustarán sobre las variables de esta muestra multivariante. En este caso, usarán como marginales distribuciones arbitrarias la Gamma, Beta y t-Student con parámetros específicos, cuyo resultado se muestra a continuación en la *Figura 4.11*, representada de forma más visual en la *Figura 4.12*:

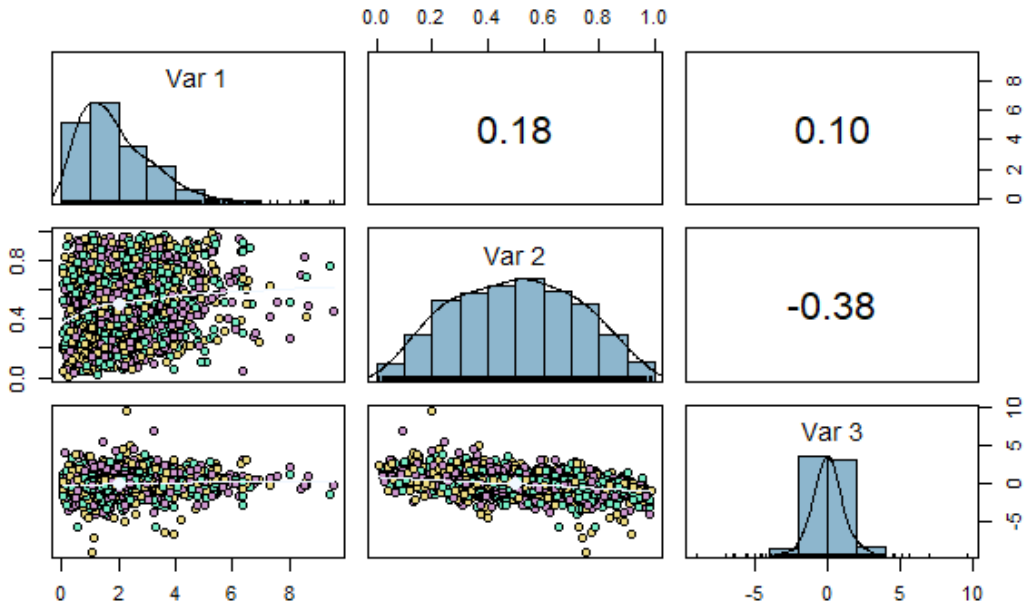

Figura 4.11 Ajuste de las distribuciones Gamma, Beta y t-Student a las variables observadas

Fuente: Elaboración propia

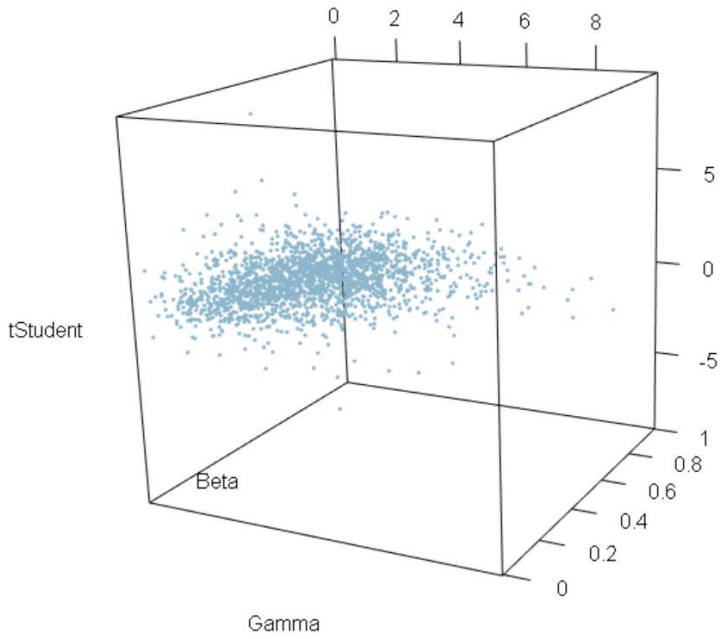

Figura 4.12 Distribuciones Gamma, Beta y t-Student

Fuente: Elaboración propia

Asimismo, para obtener la función cópula adecuada para el *set* de datos se usará la función *BiCopSelect* (Dissmann et Al., 2012), estimando sus parámetros con el método de máxima verosimilitud.

Una vez obtenido esto, se usará la función *mvdc* (Yan, 2007) para construir la función de distribución multivariante al especificar cuáles son sus marginales, aspecto que ha sido resuelto anteriormente, y la estructura de dependencia que presentan, en este caso a través de la cópula normal. Posteriormente, con la función *rMvdc* de Yan (2007) se generará la muestra aleatoria simulada a partir de esta distribución:

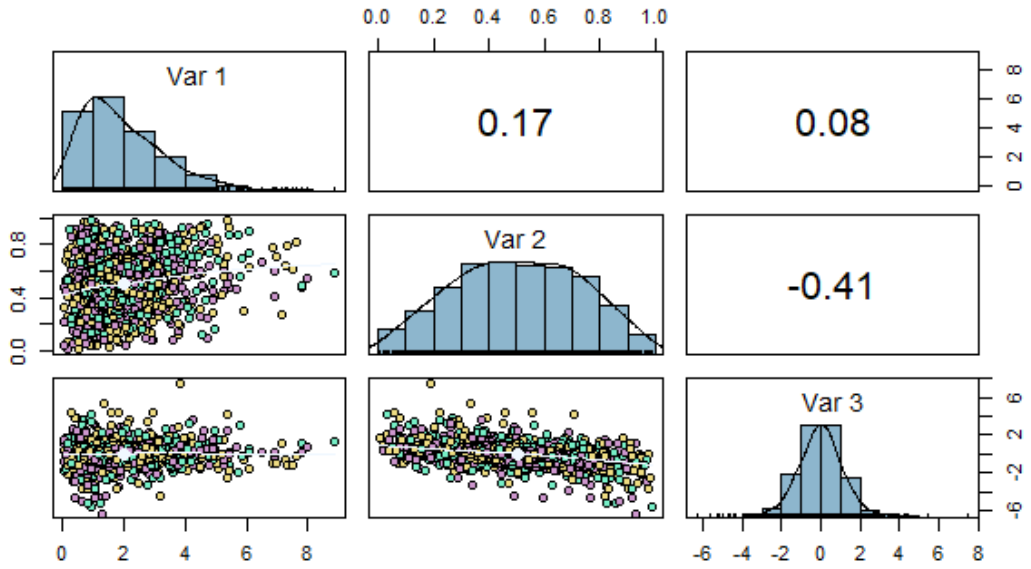

Figura 4.13 Correlaciones a pares entre las variables simuladas a través de la función cópula

Fuente: Elaboración propia basada en el estudio de Petrelli (2019)

Como se puede comprobar se ha obtenido una muestra simulada multivariante, observando en la *Figura 4.13* que se obtiene el mismo resultado que el obtenido en la *Figura 4.8*. Usando esta metodología, en el siguiente apartado se aproximará el precio de la prima de ciberseguro.

4.3. PRICING DEL CIBERSEGURO CON CÓPULAS

Hay tres elementos de riesgo que forman parte de la prima del contrato de ciberseguro *first-party* y que, por ende, serán objeto de estudio como variables aleatorias (Hearth y Herath, 2011): (1) suma asegurada, aleatoria, (2) probabilidad de ocurrencia del ciberriesgo y (3) momento del tiempo transcurrido hasta el pago de la suma asegurada. De esta forma, se denotará la variable suma asegurada como "*P*" y esta consistirá en la pérdida total en dólares por el pago de la suma asegurada (Π) en la que incurrirá una aseguradora determinada con "*q*" ordenadores. Para realizar el análisis se usarán los datos del estudio de prevalencia de ICSA (Bridwell, 2004; en Hearth y Herath, 2011) y NetDiligence (2020), donde aparecerán las pérdidas por virus y por ordenador con los datos disponibles (π). Con todo, se modelizará la distribución de la suma asegurada (Π) como una función de $\Pi = g(\pi, q)$, es decir, una función conjunta bivariante con dos distribuciones empíricas de nuestras variables aleatorias, "π" y "*q*".

Por otra parte, la ocurrencia del evento cubierto será modelada por la variable "ω", que tomará un valor entre 0 y 1 dependiendo de la probabilidad que tenga el evento de ocurrir. La tercera variable, el tiempo que transcurre hasta el pago de la suma se representará con la variable "*T*". Estas variables compondrán el valor de la prima de ciberseguro, la cual puede verse en la Ecuación (4.11).

4.3.1. Modelo de la función de distribución de la suma asegurada (Π)

En el enfoque que se usará en el presente trabajo no podrá emplear una simple regresión lineal sobre el valor de la pérdida por cada virus (π) y el número de ordenadores afectados (q), pues alguna de las dos distribuciones marginales podría no ser normal, como de hecho sucederá. De esta manera, "*q*" podría seguir cualquier distribución continua como una Pareto, una Exponencial o una Weibull (Hearth y Herath, 2011). Además, al no tratarse de marginales normales, para ver la estructura de dependencia entre ambas funciones de distribución no se podrá usar el coeficiente de correlación de Pearson (Pearson, 1895), de ahí la flexibilidad que ofrecen las cópulas.

Así, antes de nada, deberá encontrarse la cópula que mejor describa la relación de dependencia no lineal entre las dos distribuciones empíricas estudiadas, con el fin de obtener la distribución de probabilidad conjunta que permita hallar la distribución de pérdidas por la suma asegurada (Π), para la cual se propone la siguiente función (Hearth y Herath, 2011):

$$\Pi = g(\pi, q) = \begin{cases} a_1, & si\ q < l \\ a_2 + \left(\frac{q-l}{q}\right)\left(\frac{\pi}{l}\right), & si\ l \leq q < m \\ a_3 + \left(\frac{q-m}{q}\right)\left(\frac{\pi}{l}\right), & si\ q \geq m \end{cases} \quad (4.10)$$

Donde a_i, i = 1,2,3 son valores constantes y "l" y "m" son el mínimo y el máximo del número de ordenadores afectados. Como dato importante se destaca que a esta función se le aplicará la muestra bivariante obtenida con la cópula y las marginales para hallar la distribución de la suma asegurada (Π).

4.3.2. Modelo de la prima de ciberseguro a través del enfoque de cópulas

El modelo propuesto para estimar la prima neta de ciberseguro anual será el siguiente (Hearth y Herath, 2011):

$$C = \omega e^{-rT} P \quad (4.11)$$

Donde "ω" es la variable que estará entre 0 y 1, "T" es el tiempo transcurrido desde que se formaliza la póliza hasta que ocurre este evento; "r" es el tipo de interés y "P" la suma asegurada, cuya función de distribución es "Π", obtenida anteriormente y que en realidad confiere aleatoriedad al modelo. Por simplicidad, se asumirán que los eventos cubiertos solo suceden una vez y que las pérdidas provendrán de este evento único, con lo que la cobertura del contrato acabaría una vez se materializara el evento. Además, se considerará independencia entre la suma asegurada "P" y el tiempo transcurrido hasta el ciberataque "T".

Una vez introducidas las hipótesis básicas del modelo, se describirán los tres tipos de pólizas de ciberseguro cuya prima se modelizará:

i) Póliza básica *first party damage* sin franquicia:

$$\mathrm{P} = \Pi = g(\pi, \mathrm{q}) \tag{4.12}$$

ii) Póliza *first party damage* con franquicia:

$$\mathrm{P} = \begin{cases} 0, & si\ \Pi \leq \mathrm{d} \\ \Pi - \mathrm{d}, & \mathrm{si}\ \Pi > \mathrm{d} \end{cases} \tag{4.13}$$

Donde "*d*" es el umbral límite de la franquicia, como ya se vio en la Ecuación (3.4).

iii) Póliza *first party damage* con coaseguro, franquicia y límite:

$$\mathrm{P} = \begin{cases} 0, & si\ \Pi \leq d \\ (1-a)(\Pi - \mathrm{d}), & si\ d < \Pi < d + \dfrac{k}{1-a} \\ k, & si\ \Pi > d + \dfrac{k}{1-a} \end{cases} \tag{4.14}$$

En este modelo se incluye un límite de franquicia "*d*", un límite a la suma asegurada de "*k*" y un coaseguro de "*a*". Así, la aseguradora no pagará nada cuando la pérdida sea menor a "*k*", pagará el *100% (1-a)* de la pérdida y nada por encima del límite superior.

4.3.3. Tarificación del ciberseguro con datos de 2003

Para proceder al *pricing* del ciberseguro mediante cópulas, para el 2003, se usarán los datos de la encuesta ICSA (Bridwell, 2004; en Hearth y Herath, 2011). Además, se planteará el caso de una aseguradora, A, que presenta los datos reportados en la *Tabla 4.1* relativos al número de ordenadores afectados (q) y a la pérdida asociada a cada virus (π). La función de distribución bivariante que se simulará será aleatoria, puesto que el valor de las pérdidas y el número de ordenadores afectados dependerá del grado de incidencia que posea el virus y los daños que ocasione.

Además, como asume Conrad (2005), se considerará que el sistema informático de la empresa falla después de la violación, que es cuando se paga la suma asegurada, por lo que la cobertura cubriría únicamente este primer evento.

	Virus	q # de ordenadores	π $ pérdidas, en miles
1	W32/Blaster	1291	355.64
2	W32/Slammer	849	339.83
3	W32/Sobig	238	115.73
4	W32/Klez	140	65.09
5	W32/Yaha	118	45.40
6	W32/Swen	108	66.05
7	W32/Dumaru	87	39.18
8	W32/Mimail	70	19.56
9	W32/Nachi	63	20.09
10	W32/Fizzer	58	20.46
11	W32/BugBear	50	10.18
12	W32/Lirva	447	11.77
13	W32/Sober	21	6.9
14	W32/SirCam	21	5.34
15	W32/Ganda	19	7.55

Tabla 4.1 Datos de incidencias informáticas estudiadas (2003)

Fuente: Elaboración propia con datos de ICSA (Bridwell, 2004; en Hearth y Herath, 2011)

La primera cuestión a resolver al presentar estos datos es en cuanto a su estructura de dependencia: ¿realmente se puede decir que existe una relación entre ambas funciones de distribución empíricas? En caso de ser así, ¿de qué tipo? Para dar respuesta a estas preguntas se optará por la representación gráfica de ambas a través del diagrama de dispersión de la *Figura 4.14*, donde se puede comprobar que efectivamente parece que la haya.

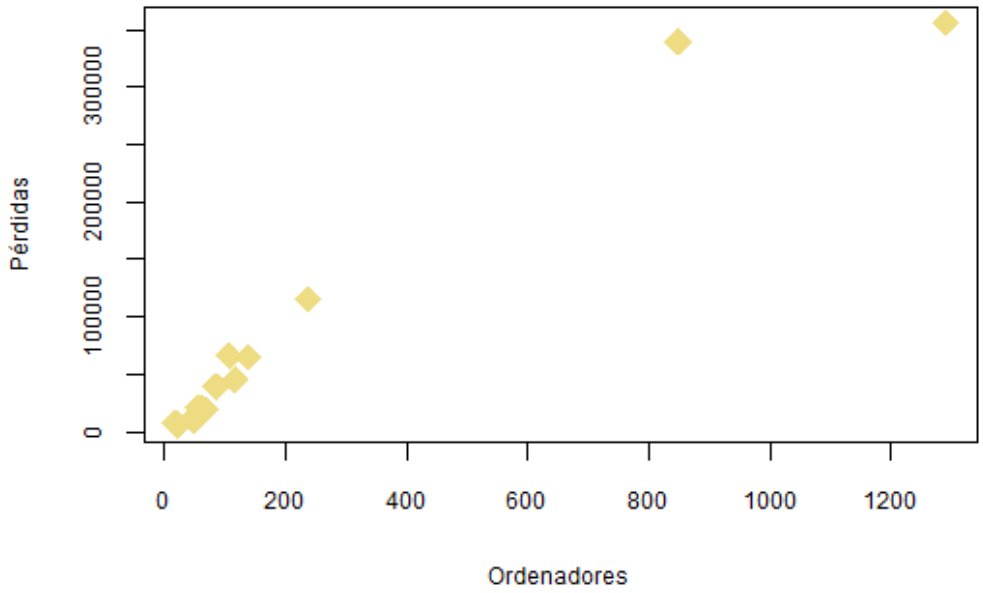

Figura 4.14 Diagrama de dispersión de las variables "q" y "π" (2003)

Fuente: Elaboración propia con datos de ICSA (Bridwell, 2004; en Hearth y Herath, 2011)

A continuación, como se explicaba anteriormente, se usará la función *BiCopSelect* (Dissmann et. Al, 2012) para encontrar la cópula que mejor se adapte a los datos, usando los estadísticos de información de Akaike (AIC) y bayesiano (BIC) (Burnham y Anderson, 2002). El resultado sería el siguiente:

"Bivariate copula: Survival Clayton (par = 10, tau = 0.83)"

Con ello, se comprueba que la cópula que mejor se adapta sería la cópula de supervivencia de Clayton representada en la *Figura 3.2*, si bien Hearth y Herath (2011) sugieren que, aunque son casi similares, la más apropiada podría sería la de Gumbel, es decir, la observada en la *Figura 3.3*.

En función del análisis realizado se elegirá la cópula estándar de Clayton (*Figura 3.1*), pues, además de que los resultados no difieren significativamente en función de la cópula escogida, las cópulas de supervivencia no permiten trabajar con distribuciones marginales, sino con distribuciones de supervivencia, aspecto que no se da; se requeriría otra herramienta más sofisticada y unos procedimientos más complejos.

Por otra parte, esta cópula en concreto se caracteriza por presentar dependencia asimétrica en las colas al tener mayor dependencia en la cola negativa que en la positiva, lo cual encaja con lo que se observaba en la *Figura 4.14*. Así, al estimar los parámetros de dicha cópula se observa que estos serían $\alpha = 10$ y $\tau = 0.83$. Con ello, es importante saber que la cópula de Clayton viene definida por la siguiente ecuación:

$$C_\alpha(u, v) = (u^{-\alpha} + v^{-a} - 1)^{-\frac{1}{\alpha}} \qquad (4.15)$$

Donde α es el parámetro de la cópula con soporte $[0, \infty)$. Al ser $\alpha > 0$, se puede aceptar que las distribuciones de ambas variables son dependientes, con dependencia positiva en la cola inferior y nula en la superior. Por otra parte, la medida no paramétrica τ de Kendall (Genest y Rivest, 1993) define el grado de dependencia entre ambas variables, siendo este elevado, y viene definida por la Ecuación (4.16):

$$\tau = \frac{\alpha}{\alpha + 2} \qquad (4.16)$$

Igualmente, Hearth y Herath (2011) indican en su ensayo que, en este caso, la distribución Weibull (Rinne, 2008) podría ser la que mejor se adaptara a ambas variables, es decir, tanto a la pérdida como al número de ordenadores afectados, ya que unos pocos virus (entre el 15% y el 25%) afectan a un intervalo entre el 75% y el 85% de los ordenadores. Así, se intentarán probar alternativas a sus resultados y se tomarán decisiones sobre las mismas ajustando varias distribuciones continuas a las variables gracias a la función *fitdist* (Cullen y Frey, 1999):

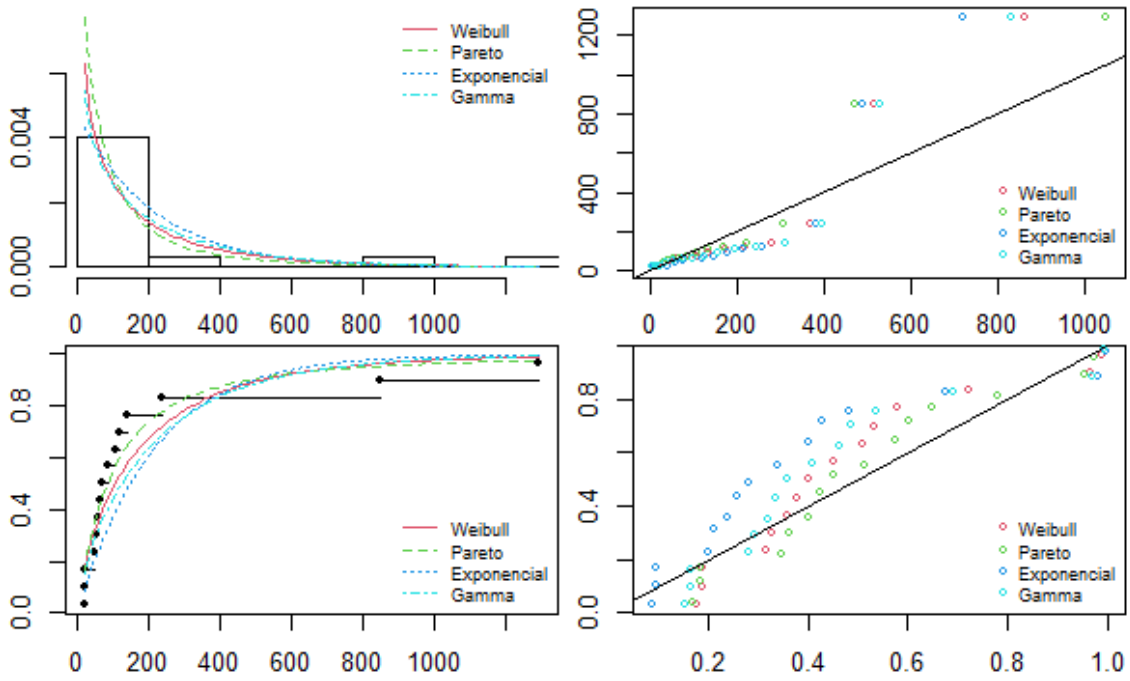

Figura 4.15 Ajuste de distribuciones continuas a la variable "q", número de ordenadores (2003). De izquierda a derecha y de arriba abajo: histograma y densidades teóricas, Q-Q *plot*, CDF teóricas, P-P *plot*

Fuente: Elaboración propia

Como se puede observar en la *Figura 4.15*, a diferencia de lo expuesto por Hearth y Herath (2011) la distribución que mejor parece ajustarse sería la Pareto (Arnold y Laguna, 1977), sobre todo en la cola inferior. Con el fin de comprobar esto, se emplearán los criterios AIC y BIC y los test de Kolmogorov-Smirnov (Kolmogorov, 1933), Cramer-von Mises (Cramér y von Mises, 1928) y Anderson-Darling (Anderson y Darling, 1952), cuyos resultados se exponen a continuación en la *Tabla 4.2* y *Tabla 4.3* respectivamente:

	Exponencial	Gamma	Pareto	Weibull
Anderson-Darling	2.16	1.29	0.59	0.98
Cramer-von Mises	0.41	0.24	0.09	0.16
Kolmogorov-Smirnov	0.32	0.27	0.17	0.22

Tabla 4.2 Estadísticos de bondad del ajuste de la variable "q" (2003)

Fuente: Elaboración propia

	Exponencial	Gamma	Pareto	Weibull
Criterio de Akaike (AIC)	192.7	193.3	188.83	191.87
Criterio Bayesiano (BIC)	193.41	194.71	190.25	193.28

Tabla 4.3 Criterios de bondad del ajuste de la variable "q" (2003)

Fuente: Elaboración propia

En base a los cálculos realizados se confirma que, efectivamente, la distribución que mejor se ajusta a los datos de la variable "número de ordenadores" no será la Weibull sino la Pareto, pues arroja valores inferiores tanto en los estadísticos como en los criterios. Por otro lado, realizando el mismo procedimiento, se observa que para la variable "pérdidas" los resultados de ajustar las diferentes distribuciones serían:

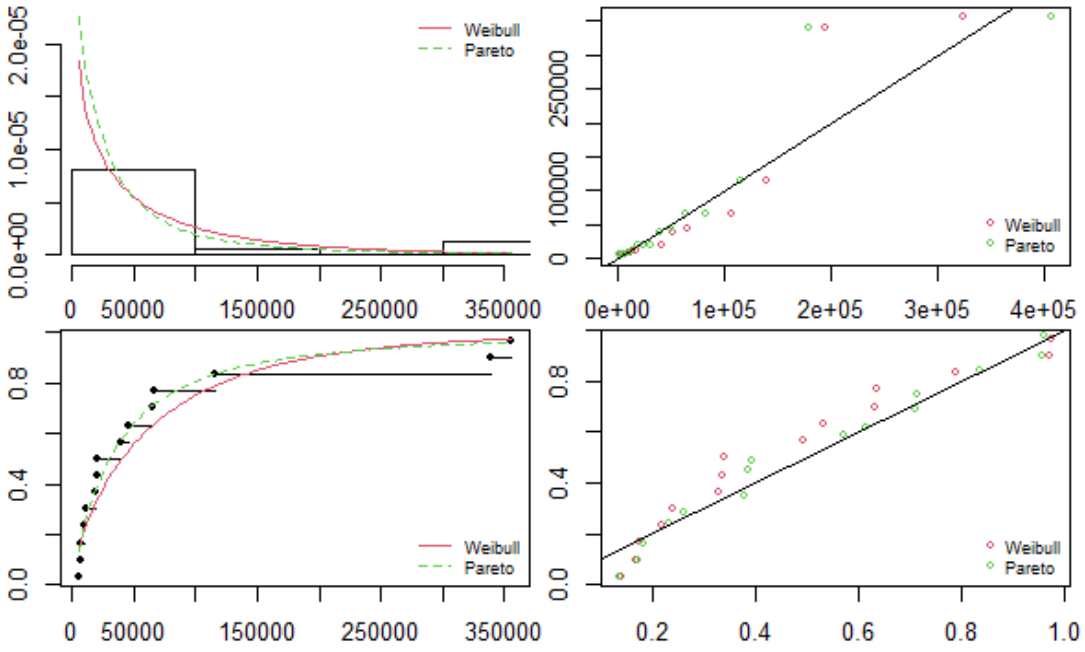

Figura 4.16 Ajuste de distribuciones continuas a la variable "π", pérdidas (2003). De izquierda a derecha y de arriba abajo: histograma y densidades teóricas, Q-Q *plot*, CDF teóricas, P-P *plot*

Fuente: Elaboración propia

Destacando que, en este caso, existe una limitación al omitirse las distribuciones Exponencial y Gamma, en vista que, al tratarse de valores elevados, R los eleva a la exponencial y se disparan al infinito, llevando ello a error. A pesar de ello, en la *Figura 4.16* se comprueba que de nuevo la Pareto parece ajustarse de mejor forma a esta segunda variable, lo cual se comprueba a continuación con los mismos estadísticos y criterios usados en la *Tabla 4.2* y *Tabla 4.3*:

	Pareto	Weibull
Anderson-Darling	0.36	0.68
Cramer-von Mises	0.04	0.11
Kolmogorov-Smirnov	0.14	0.19

Tabla 4.4 Estadísticos de bondad del ajuste de la variable "π" (2003)

Fuente: Elaboración propia

	Pareto	Weibull
Criterio de Akaike (AIC)	366.54	368.33
Criterio Bayesiano (BIC)	367.96	369.74

Tabla 4.5 Criterios de bondad del ajuste de la variable "π" (2003)

Fuente: Elaboración propia

De nuevo, como conclusión de los resultados expuestos en la *Tabla 4.4* y la *Tabla 4.5*, la distribución Pareto será la que proporcione un mayor grado de ajuste a la variable "pérdidas". Con todo, una vez visto que ambas variables siguen esta distribución, sus parámetros se recogen en la *Tabla 4.6*:

	Ordenadores (q) - Pareto	Pérdidas (π) - Pareto
Shape	1.70	1.62
Scale	164.69	57,035.47

Tabla 4.6 Parámetros de la distribución ajustada para ambas variables (2003)

Fuente: Elaboración propia

En términos generales, sabiendo que la cópula seleccionada fue la de Clayton, con parámetros $\alpha = 10$ y $\tau = 0.83$, y que las marginales de ambas variables seguirán una Pareto, cuyos parámetros se recogen en la *Tabla 4.6*, se procederá a construir la distribución conjunta bivariante con la función *mvdc* (Yan, 2007) indicando la estructura de dependencia a través de la cópula y marginales mencionadas. Posteriormente, con la función *rMvdc* de Yan (2007) se generará una muestra aleatoria simulada que será contrastada con los valores observados:

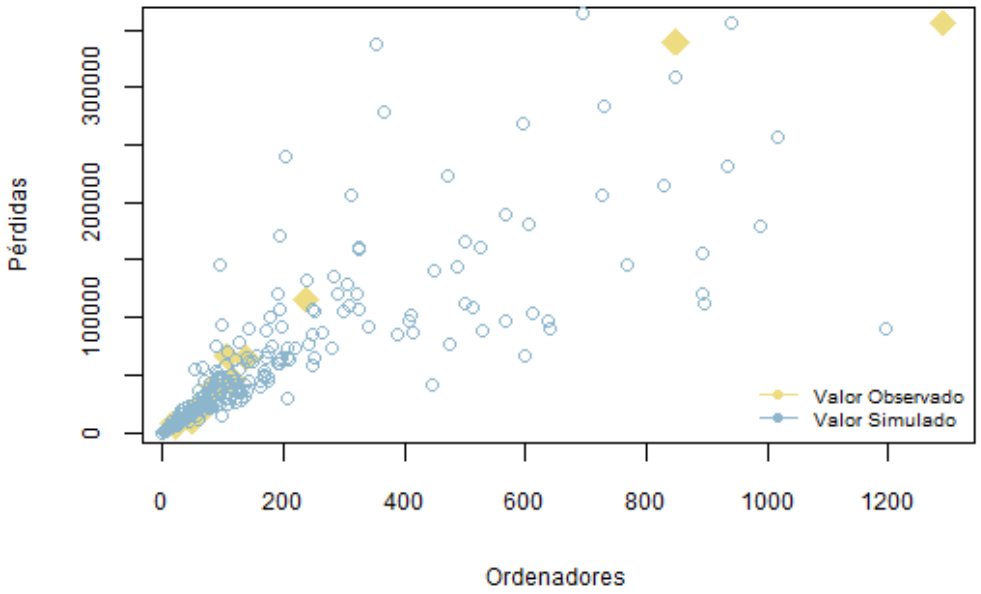

Figura 4.17 Valores observados y simulados a través de la cópula de Clayton, con marginales Pareto (2003)

Fuente: Elaboración propia

Con el objetivo de verificar que la muestra bivariante simulada arroja la misma correlación entre las variables que la existente entre las originales, se ha recurrido al coeficiente de correlación de Spearman, cuyos resultados se recogen en la *Tabla 4.7*, donde se observa que son casi idénticos. En base a ello, se da por adecuado el procedimiento realizado.

	Muestra Observada	Muestra Simulada
Coeficiente de correlación de Spearman	0.9598	0.9564

Tabla 4.7 Comparación del coeficiente de correlación de Spearman muestra observada vs. Simulada (2003)

Fuente: Elaboración propia

El siguiente paso será usar la muestra aleatoria obtenida en la Ecuación (4.10) para hallar la distribución de pérdidas por el pago de la suma asegurada. Además, se considera que $a_1 = 400$, $a_2 = 125$, $a_3 = 300$ (Hearth y Herath, 2011), con límites correspondientes el mínimo y el máximo de ordenadores infectados, en este caso 19 y 1,291. La densidad de esta distribución obtenida puede ser observada en la *Figura 4.18*:

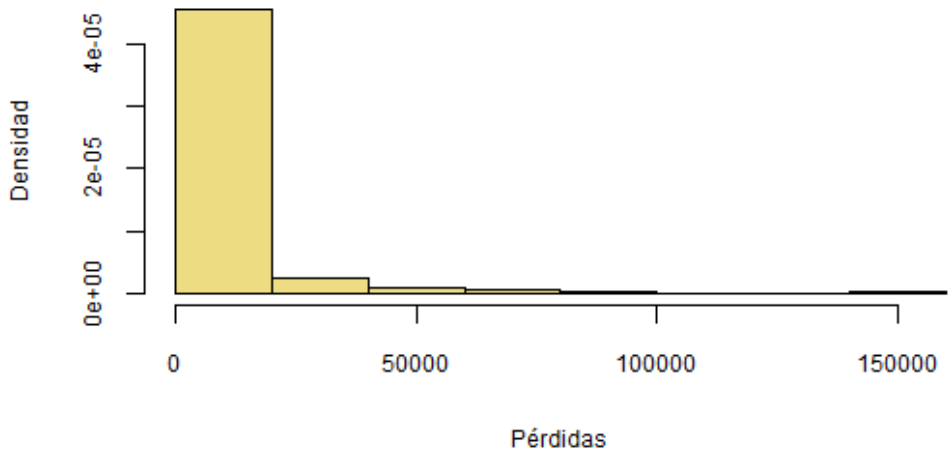

Figura 4.18 Histograma función de distribución de pérdidas (2003)

Fuente: Elaboración propia

Por último, a través de estos datos se calculará la prima de seguro según las diferentes pólizas vistas anteriormente en las Ecuaciones (4.12), (4.13) y (4.14), sabiendo que la prima viene definida por la Ecuación (4.11).

Donde:

- ω es una variable con probabilidad 0.052 (United States Census Bureau, 2008; en Petrelli, 2019).
- r es el tipo de interés, estimado en función del tipo de la Reserva Federal de USA de marzo de ese año, un 1.25% (Expansión, 2020).
- T es el tiempo en días hasta que acaece la violación de seguridad, estimado en 197 días (Ponemon Institute LLC, 2018; en Petrelli, 2019).
- P es la suma asegurada a pagar al cliente en caso de incurrir el evento asegurado, que cuya función de distribución será obtenida a partir de simulación de Monte Carlo.

a) Póliza básica *first party damage* sin franquicia: en este tipo de póliza aplicarán directamente las fórmulas empleadas en la Ecuación (4.11) y (4.12), pues $P = \Pi = g(\pi, q)$. En función de esto, para una compañía con un rango de ordenadores afectados entre 19 y 1,291 se ha simulado una prima anual de ciberseguro de $400. Este dato se obtendrá al hacer una media de 10,000 simulaciones de las posibles primas para esta póliza.

b) Póliza *first party damage* con franquicia: en este caso, al no cumplirse la igualdad anterior se aplica la Ecuación (4.13) para hallar la función de distribución de pérdidas. Posteriormente, se emplea la Ecuación (4.11) para hallar la prima. Con el objetivo de transmitir al lector la influencia de la franquicia en el cálculo de la prima, en la *Tabla 4.8* se ilustrará cómo varía la misma según el importe franquiciado:

Franquicia, "d"	Prima ($)
0	400
500	376
1,000	357
1,500	340
2,000	324
2,500	309
3,000	297
3,500	285
4,000	274

Tabla 4.8 Variación de la prima de ciberseguro en función al importe de la franquicia (2003)

Fuente: Elaboración propia

Con dichos resultados se obtiene que, al elevar el importe de la franquicia, la prima por ordenador irá descendiendo, puesto que la aseguradora no pagará todo el siniestro, sino a partir de una cantidad en concreto ("d"). Así, a medida que aumenta esta cantidad reasegurada el riesgo asumido por la compañía será cada vez menor. A la vez, se puede comprobar que, si la franquicia fuera de 0, el importe coincidiría con la prima calculada para la primera póliza (a).

c) Póliza *first party damage* con coaseguro y límite:

Para hallar la función de distribución de pérdidas, al no cumplirse de nuevo la igualdad del primer apartado, se aplica la Ecuación (4.14), donde se irán variando los valores del límite de la suma asegurada, "k", la porción de riesgo coasegurada, "α", y la franquicia, "d", cuyos resultados se recogen en la *Tabla 4.9*:

Límite asegurado (k)	Coaseguro (α)	Franquicia (d)	0	500	1,000	1,500	2,000	2,500
20,000	5%	Prima	262	241	225	210	197	185
	10%		253	232	217	203	190	179
	15%		243	224	209	196	184	173
	20%		233	215	201	188	177	167
15,000	5%		238	218	202	188	175	164
	10%		230	210	195	182	170	159
	15%		221	203	189	176	164	154
	20%		213	195	182	170	158	149
10,000	5%		204	185	171	158	146	136
	10%		198	179	166	153	142	132
	15%		191	173	160	148	138	128
	20%		184	167	155	143	133	124

Tabla 4.9 Variación de la prima de ciberseguro en función al importe de la franquicia, del coaseguro y del límite asegurado (2003)

Fuente: Elaboración propia

En este caso, recordando que la aseguradora no pagaba nada por debajo de la franquicia ("d"), "(1- α)" por encima de la misma, y nada por encima del límite ("k"), se obtienen las siguientes conclusiones:

- A mayor tasa de coaseguro (α) la compañía soporta una menor porción de ciberriesgo, por lo que las primas van disminuyendo.
- En cuanto al límite en la suma asegurada (k) cuando este se reduce, al disminuir en consecuencia el importe que la aseguradora tendría que pagar, resulta en un decremento de la prima. De esta manera, manteniendo la tasa de coaseguro constante en un 5% y nada de franquicia (d = 0), se observa que la prima disminuye de $262 a $238 y $204 si dicho límite baja de $20,000 a $15,000 y $10,000 respectivamente.
- A mayor franquicia menor serán las primas, al asumir la aseguradora una menor porción de riesgo.

4.3.4. Tarificación del ciberseguro con datos de 2020

Para los datos de 2020, se hallará la prima de ciberseguro para otra aseguradora, B, con el mismo procedimiento, pero usando los datos obtenidos del reporte de NetDiligence (2020), que en este caso no reporta los virus en concreto que afectan a los ordenadores sino la causa en general, definiendo el número de ordenadores afectados y el total de pérdidas en la *Tabla 4.10*:

	Causa de la pérdida	q # de ordenadores	π $ pérdidas, en miles
1	Email de trabajo infectado	318	41.9
2	Hacker	430	272.2
3	Hacker/Malware/Virus combinados	606	299
4	Acciones legales	56	7.2
5	Pérdida/Ordenadores robados/combinados	127	6.4
6	Ordenadores robados	100	5
7	Malware/Virus	176	26.3
8	Negligencia	5	0.389
9	Paper Records	31	1.5
10	Phishing	258	17.8
11	Error de programación	20	7.1
12	Ransomware	915	130.5
13	Problemas con empleados	161	12.3
14	Ingeniería social combinada	782	79.4
15	Ingeniería social (tipo desconocido)	157	11.2
16	Errores del staff	344	5.7
17	Fallos del sistema	12	19.4
18	Robo de dinero	10	0.381

19	Acciones legales terceras partidas	15	0.4
20	Trademark/Copyright	9	1.3
21	Fraude en transferencias bancarias	191	26.8
22	Recogida de datos errónea	2	0.091

Tabla 4.10 Datos de incidencias informáticas estudiadas (2020)

Fuente: Elaboración propia con datos de NetDiligence (2020)

Al igual que ocurría con los datos de 2003, la primera cuestión a resolver será si existe relación entre ambas distribuciones empíricas y de qué tipo en caso afirmativo. Para verlo, la *Figura 4.19* recoge el gráfico de dispersión correspondiente, donde también parece que la exista, usando estos datos como una muestra representativa ajustada para el volumen de la compañía anterior, lo cual permitirá realizar una comparación directa entre ambas:

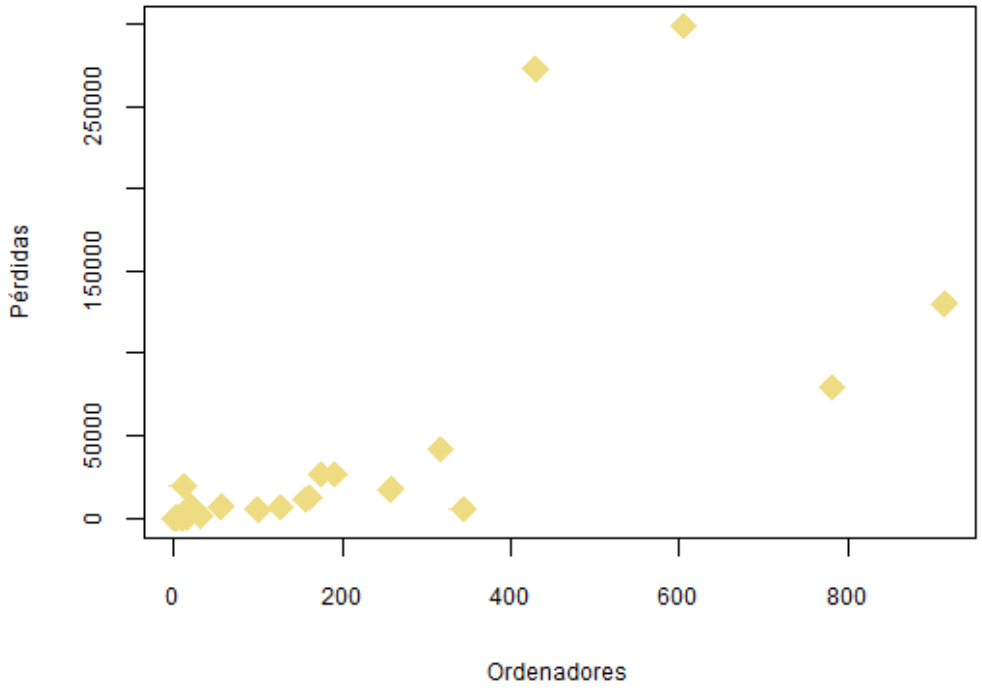

Figura 4.19 Diagrama de dispersión de las variables "q" y "π" (2020)

Fuente: Elaboración propia

El siguiente paso será encontrar la cópula que recoja mejor la relación entre los datos, en función de los estadísticos AIC y BIC, siendo el resultado:

"Bivariate copula: Survival Gumbel (par = 3.38, tau = 0.7)"

Siguiendo el mismo razonamiento que se ha usado anteriormente nos valdremos del consejo de la función, pero por simplicidad en los cálculos se usará la cópula de Gumbel, que presenta la siguiente forma:

$$C_\alpha(u,v) = \exp\left\{-[((-\ln u)^\alpha + (-\ln v)^\alpha]^{\frac{1}{\alpha}}\right\} \qquad (4.17)$$

Donde α es el parámetro de la cópula con soporte [1, ∞). Al ser α diferente de la unidad, se puede comprobar que las distribuciones de ambas variables son también dependientes. Por otra parte, la medida τ de Kendall estaba definida por la Ecuación (4.16), dando como resultado los parámetros de la cópula: α = 3.38 y τ = 0.7.

En cuanto a la distribución que mejor se adapta a cada una de las variables, de nuevo se recurrirá a la función *fitdist* (Cullen y Frey, 1999) con el fin de observar gráficamente el grado de ajuste:

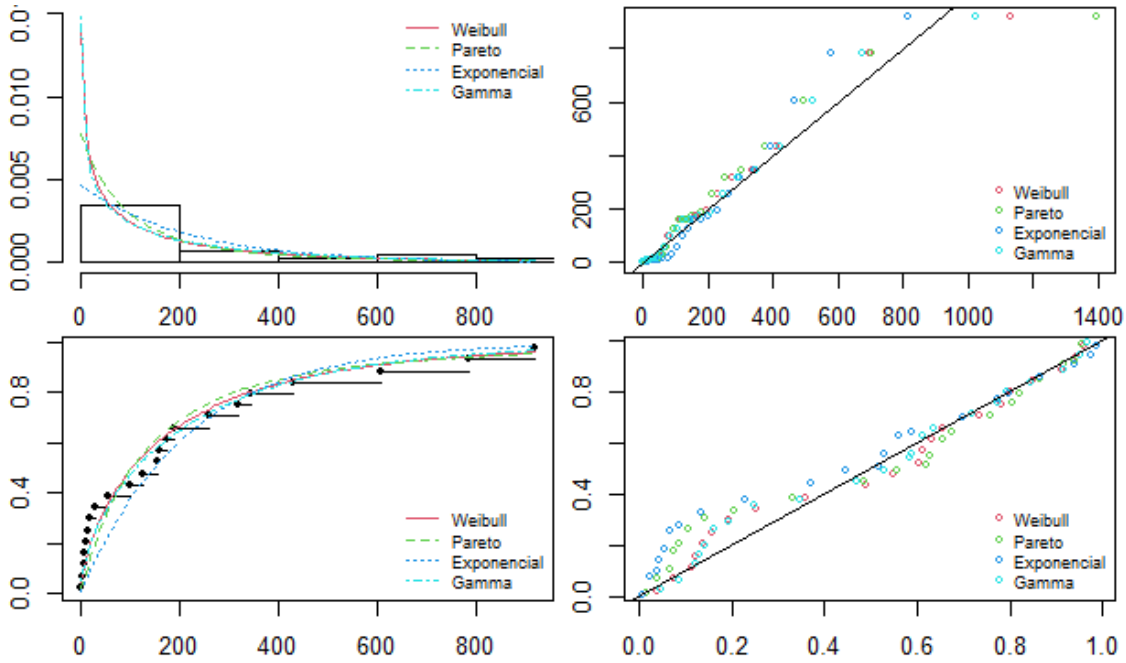

Figura 4.20 Ajuste de distribuciones continuas a la variable "q", número de ordenadores (2020). De izquierda a derecha y de arriba abajo: histograma y densidades teóricas, Q-Q *plot*, CDF teóricas, P-P *plot*

Fuente: Elaboración propia

A simple vista, en la *Figura 4.20* se observa que la mejor distribución será o bien la Weibull o bien la Gamma. Con ello, para elegir la distribución más adecuada para el número de ordenadores, con datos de 2020, se emplearán los estadísticos y criterios pertinentes, cuyos resultados se exponen en la *Tabla 4.11* y la *Tabla 4.12*:

	Exponencial	Gamma	Pareto	Weibull
Anderson-Darling	1.89	0.32	0.89	0.35
Cramer-von Mises	0.21	0.05	0.13	0.06
Kolmogorov-Smirnov	0.23	0.12	0.18	0.13

Tabla 4.11 Estadísticos de bondad del ajuste de la variable "q" (2020)

Fuente: Elaboración propia

	Exponencial	Gamma	Pareto	Weibull
Criterio de Akaike (AIC)	282.26	279.62	282.56	279.76
Criterio Bayesiano (BIC)	283.35	281.80	284.74	281.94

Tabla 4.12 Criterios de bondad del ajuste de la variable "q" (2020)

Fuente: Elaboración propia

Indudablemente, los resultados estimados en la *Tabla 4.11* y la *Tabla 4.12* indican que para la variable del número de ordenadores cabría ajustar una Gamma, si bien la Weibull tampoco arrojaría resultados muy diferentes. Por otro lado, para la variable "pérdidas":

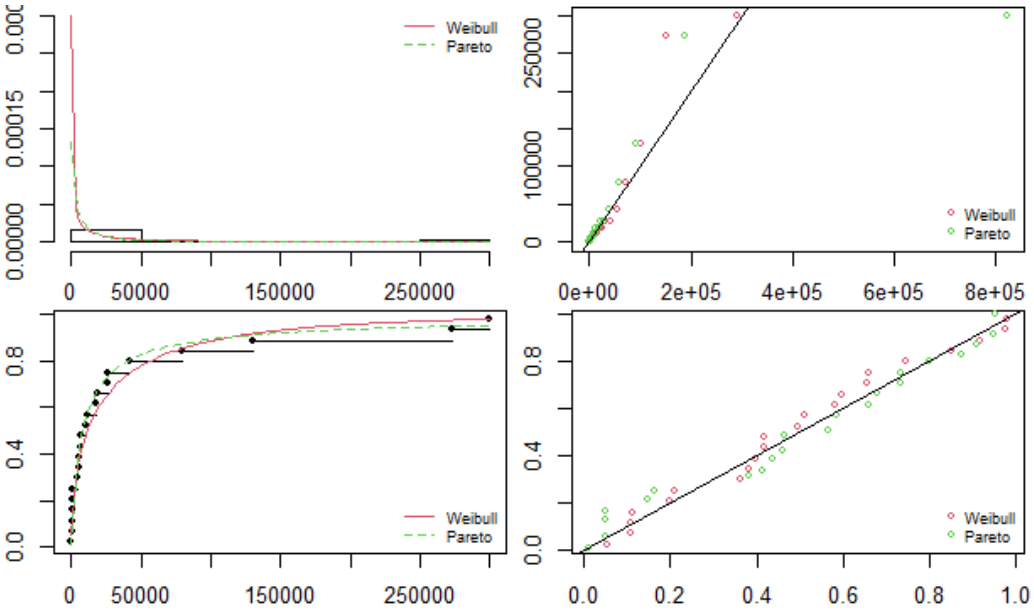

Figura 4.21 Ajuste de distribuciones continuas a la variable "π", pérdidas (2020). De izquierda a derecha y de arriba abajo: histograma y densidades teóricas, Q-Q *plot*, CDF teóricas, P-P *plot*

Fuente: Elaboración propia

Como sucedía con los datos de 2003, debe tenerse en cuenta con la limitación de no poder aplicar las distribuciones Exponencial y Gamma. Aun así, gracias a la *Figura 4.21* puede verse que, en este caso, la Weibull parece ajustarse mejor a la variable de pérdidas, aspecto que puede ser contrastado con los métodos que se han ido siguiendo en la *Tabla 4.13* y *Tabla 4.14*:

	Pareto	Weibull
Anderson-Darling	0.49	0.31
Cramer-von Mises	0.06	0.04
Kolmogorov-Smirnov	0.13	0.11

Tabla 4.13 Estadísticos de bondad del ajuste de la variable "π" (2020)

Fuente: Elaboración propia

	Pareto	Weibull
Criterio de Akaike (AIC)	498.05	496.66
Criterio Bayesiano (BIC)	500.24	498.84

Tabla 4.14 Criterios de bondad del ajuste de la variable "π" (2020)

Fuente: Elaboración propia

Una vez que se dispone de esta información, se resumirán los parámetros que siguen las distribuciones de ambas variables en la *Tabla 4.15*:

	Ordenadores (q)	Pérdidas (π)
Shape	0.6	0.52
Rate/Scale	0.003	23,044.91

Tabla 4.15 Parámetros de la distribución ajustada para ambas variables (2020)

Fuente: Elaboración propia

Así, una vez visto que la cópula seleccionada en este caso ha sido la de Gumbel, con parámetros $\alpha = 3.38$ y $\tau = 0.7$, y que las marginales correspondientes a las variables serían una Gamma y una Weibull respectivamente, con los parámetros recogidos en la Tabla *4.15*, se procederá a construir la distribución multivariante y la muestra aleatoria bivariante que será contrastada con los valores observados:

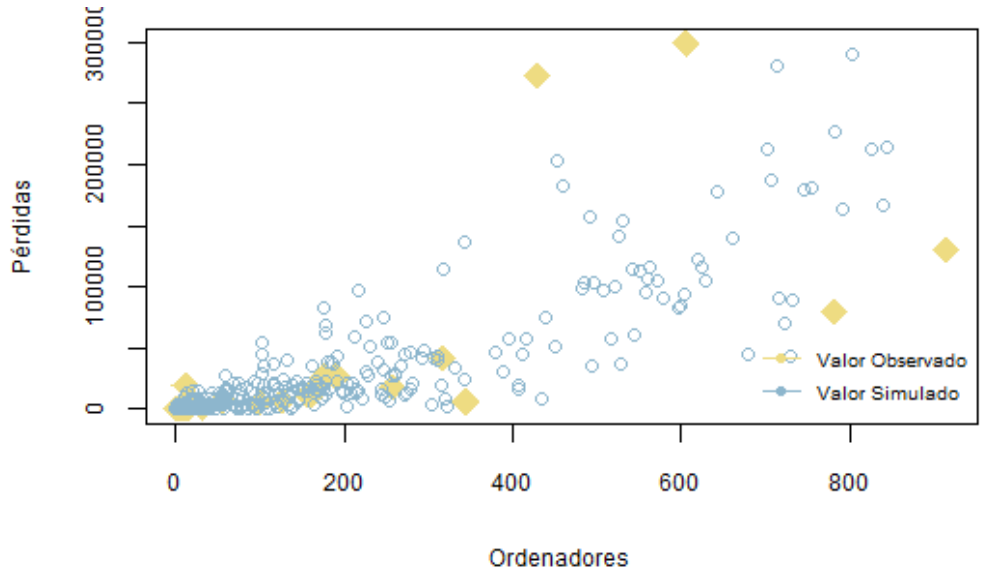

Figura 4.22 Valores observados y simulados a través de la cópula de Gumbel, con marginales Gamma y Weibull

Fuente: Elaboración propia

Para comprobar que la muestra simulada tiene la misma correlación que la original, como se ha realizado anteriormente se aplica el coeficiente de correlación de Spearman, con los resultados obtenidos en la *Tabla 4.16*. Con ello, se considerará que el proceso realizado es válido:

	Muestra Observada	Muestra Simulada
Coeficiente de correlación de Spearman	0.8464	0.8646

Tabla 4.16 Comparación del coeficiente de correlación de Spearman muestra observada vs. Simulada (2020)

Fuente: Elaboración propia

Con todo, con la muestra obtenida se aplicará simulación de Monte Carlo para hallar la distribución de pérdidas de la suma asegurada, con los mismos supuestos planteados con los datos de 2003, con lo que resulta la densidad de la distribución obtenida en la Figura *4.23*:

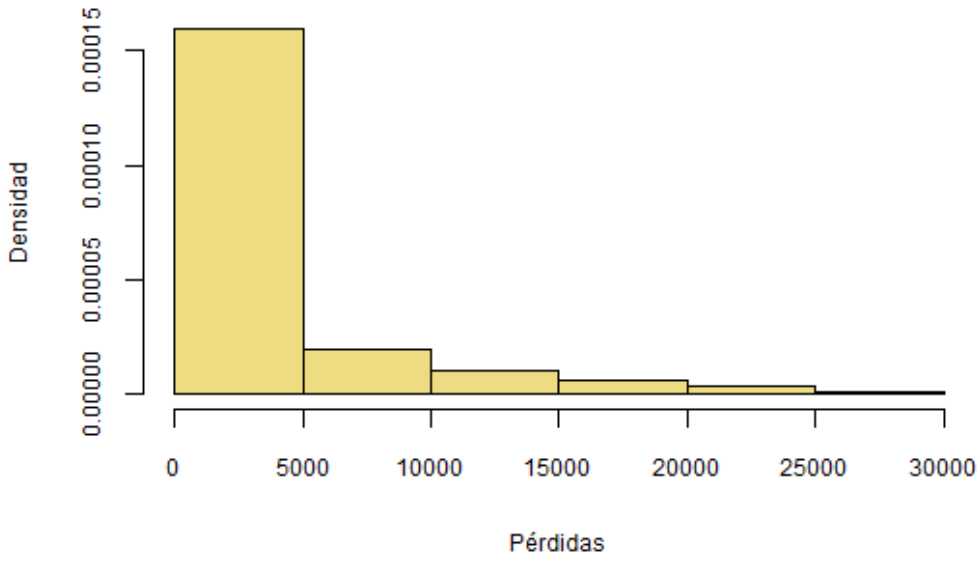

Figura 4.23 Histograma función de distribución de pérdidas (2020)

Fuente: Elaboración propia

El último paso será calcular la prima de seguro según las diferentes pólizas, modificando las variables que definen la prima según la Ecuación (4.11), donde:

- ω es una variable con probabilidad 0.25 (NetDiligence, 2020).
- r es el tipo de interés, estimado en función del tipo de la Reserva Federal de USA de marzo de 2021, el cual ha caído al 0% (Expansión, 2020).
- T es el tiempo en días hasta que acaece la violación de seguridad, estimado en 207 días (Sobers, 2021).
- P es la suma asegurada a pagar al cliente en caso de incurrir el evento asegurado, cuya distribución será obtenida a partir de la simulación.

a) Póliza básica *first party damage* sin franquicia: en este tipo de póliza se usarán directamente las fórmulas empleadas en la Ecuación (4.11) y (4.12), ya que $\mathrm{P} = \Pi = g(\pi, \mathrm{q})$, resultando una prima anual de ciberseguro promedio de $864 en caso de que la compañía asegurara entre 2 y 915 ordenadores, lo cual representaría un aumento de un 115.87% con respecto a 2003.

   El motivo de esta subida se debe, principalmente, al aumento de la probabilidad de sufrir un ciberataque, más pronunciado aún si cabe en los últimos años, como se han ido estudiando a lo largo del trabajo.

b) Póliza *first party damage* con franquicia:

Franquicia, "d"	Prima ($)
0	864 (+115.87%)
500	757 (+101.5%)
1,000	679 (+90.39%)
1,500	613 (+80.47%)
2,000	559 (+72.66%)
2,500	514 (+66.25%)
3,000	474 (+59.97%)
3,500	437 (+53.51%)
4,000	403 (+47.09%)

Tabla 4.17 Variación de la prima de ciberseguro en función al importe de la franquicia (2020)

Fuente: Elaboración propia

En este caso se observa que, a mayor importe de la franquicia, menor será la variación de las primas de 2020 con respecto a las de 2003, llegando a estabilizarse esta diferencia en los mayores umbrales. Además, al aumentar el importe de la franquicia decrecerá la prima por ordenador, ya que el riesgo asumido por la compañía será cada vez menor.

De esta forma, se llega a la conclusión de que el aumento promedio con respecto al 2003 ha sido de un 76.41% en este tipo de póliza.

c) Póliza *first party damage* con coaseguro, franquicia y límite:

Límite asegurado (k)	Coaseguro (α)	Franquicia (d)	0	500	1,000	1,500	2,000	2,500
20,000	5%	Prima	807	708	635	573	522	481
	10%		768	673	603	544	496	457
	15%		727	637	572	516	471	434
	20%		687	602	540	487	445	410
15,000	5%		767	672	603	544	498	459
	10%		736	645	579	523	477	440
	15%		704	617	553	499	456	421
	20%		669	586	526	475	434	401
10,000	5%		685	594	530	475	432	398
	10%		660	574	512	459	418	386
	15%		635	552	493	443	404	372
	20%		608	529	473	426	388	358

Tabla 4.18 Variación de la prima de ciberseguro en función al importe de la franquicia, del coaseguro y del límite asegurado (2020)

Fuente: Elaboración propia

En este último supuesto, al añadir el límite a la suma asegurada y el coaseguro a la franquicia, en término medio las primas también aumentan con respecto a las de 2003, en particular un 191.90%. Además, el importe mínimo de la prima es de $358, correspondiente a un límite (k) de $10,000, un coaseguro del 20% y una franquicia de $2,500, es decir, a mayor coaseguro, mayor franquicia y menor límite menor será la prima. Por otra parte, la prima máxima será de $807, correspondientes a un 5% de coaseguro y a un límite en la suma asegurada de $20,000, sin franquicia.

5. CONCLUSIONES

En este Trabajo de Fin de Máster se ha pretendido en todo momento perseguir un objetivo principal y uno secundario. Para último se ha realizado un estudio teórico sobre la problemática que plantea el ciberriesgo y la estructuración del sector en general, así como de sus elementos principales, con el fin de poder entender el objetivo principal: la propuesta de una metodología estadístico-actuarial sobre la modelización de la prima de ciberseguro a través de las funciones cópula y la simulación de Monte Carlo.

Con ello, se observa como la estructura de dependencia de los datos de 2003 podía modelizarse con la cópula de Clayton y marginales Pareto, siendo los de 2020 modelados con la cópula de Gumbel y marginales Gamma y Weibull, experimentando aumentos en la prima de ciberseguro del 115% entre ambos periodos, porcentaje que es susceptible de variación al añadir diferentes elementos de reaseguro como el coaseguro, el límite a la suma asegurada y la franquicia, toda vez que se ha demostrado que estos contribuyen de forma significativa a la mitigación del riesgo en las aseguradoras.

En conclusión, puede considerarse que se ha cumplido el objetivo principal al poder aplicar la metodología planteada para estimar las primas en ambos periodos y para diferentes pólizas, realizando previamente un planteamiento teórico del contexto relativo al ciberseguro. Como opinión personal subrayar que, a través de este trabajo, se ha intentado buscar un razonamiento que incluyera todos los elementos posibles que pudieran afectar al proceso de tarificación del ciberseguro, proponiendo para ello un procedimiento que podría solventar el problema de escasez de modelos cuantitativos en el sector. Este permitiría que, a medida que se fueran disponiendo de datos referentes a ciberriesgos, se fuera cambiando la modelización de los productos, lo cual aportaría flexibilidad y disminuiría la disparidad que produce en las primas el que cada compañía realice un modelo propio ad-hoc.

Igual de importante es también observar que estamos viviendo actualmente un periodo de plena transformación digital con constantes cambios, lo cual podría alterar la probabilidad de ser afectados, en mayor o menor medida, por un ciberataque, lo cual haría que el precio de la prima se incrementara o decrementara a su vez. De ahí la importancia de considerar la "volatilidad" de este escenario futuro para determinar finalmente la prima de las ciberpólizas con la mayor precisión posible, al recoger todos los factores que puedan influenciar a la severidad del mismo.

Por aspectos como este, y citando a Laurence Peter, un pedagogo canadiense: "un [actuario] es un experto que sabrá mañana por qué las cosas que predijo ayer no han sucedido hoy".

6. BIBLIOGRAFÍA

Agencia Española de Protección de Datos. (2019). *Principios RGPD*. Disponible en línea en: https://www.aepd.es/es/derechos-y-deberes/cumple-tus-deberes/principios#:%7E:text=Principio%20de%20%E2%80%9C%20licitud%2C%20transparencia%20y,y%20transparente%20para%20el%20interesado.&text=Una%20vez%20que%20esas%20finalidades,permita%20identificar%20a%20los%20interesados [Consultado 11-03-2021]

Alonso, R. (2021). *Los centros sanitarios, en peligro por el aumento de los ciberataques y el despliegue de la nube*. ABC. https://www.abc.es/tecnologia/redes/abci-centros-sanitarios-peligro-aumento-ciberataques-y-despliegue-nube-202102191426_noticia.html [Consultado 23-04-2021]

Anderson, T. W. y Darling, D. A. (1952). "Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic processes". Annals of Mathematical Statistics. 23: 193–212. doi:10.1214/aoms/1177729437.

Arnold, B.C. y Laguna, L. (1977). On Generalized Pareto Distributions with Applications to Income Data, International Studies in Economics Monograph No. 10, Department of Economics, Iowa State University, Ames, IA.

Arsys. (2019). *Qué son las herramientas EDR y cómo mejoran la seguridad*. Disponible en línea en: https://www.arsys.es/blog/edr-seguridad/ [Consultado 09-02-2021]

Bichara, D., Kang, Y., Castillo-Chavez, C., Horan, R. y Perrings, C. (2015). SIS and SIR epidemic models under virtual dispersal, Bull Math Biol., Vol. 77, No. 11, pp. 2004-2034

Bridwell, L. (2004). "ICSA Labs 9th Annual Computer Virus Prevalence Survey". ICSA Labs. Disponible en línea en: https://www.icsalabs.com/icsa/docs/html/library/whitepapers/VPS2003.pdf [Consultado 09-03-2021]

Burnham, K. P., y Anderson, D.R. (2002). Model selection and multimodel inference, a practical information-theoretic approach. Second edition. Springer, New York, New York, USA.

CCDCOE (2012). *Tallinn Manual on international law applicable to cyber warfare*. Nato CCDCOE, Rule 3091. Disponible en línea en http://csef.ru/media/articles/3990/3990.pdf [Consultado 21-03-2021]

Cintas del Río, R. (2007). *Teoría de cópulas y control de riesgo financiero*. Universidad Complutense de Madrid. Disponible en línea en: https://www.ucm.es/data/cont/media/www/pag-41459/Copulas.pdf [Consultado 10-05-2021]

Clayton, D. G. (1978). A Model for Association in Bivariate Life Tables and Its Application in Epidemiological Studies of Familial Tendency in Chronic Disease Incidence, Biometrika, Vol. 6, No. 1, pp. 141-151;

Conrad, J.R. (2005). Analyzing the risks of information security investments with Monte-Carlo simulations, Workshop on the Economics of Information Security (WEIS). Harvard University.

Cramér, H. (1928). "On the Composition of Elementary Errors". Scandinavian Actuarial Journal. 1928 (1): 13–74. doi:10.1080/03461238.1928.10416862.

Cullen, A. y Frey, H. (1999). Probabilistic Techniques in Exposure Assessment. Plenum Publishing Co., 1st edition.

De las Casas, J. (2019). *La prevención de ciberriesgos, una asignatura pendiente para las pymes*. Expansión. Disponible en línea en: https://www.expansion.com/empresas/2019/07/23/5d374126e5fdea31118b460b.html [Consultado 08-03-2021]

Deloitte. (2017). *Hacklab WannaCry. Análisis del ataque y lecciones aprendidas*. Deloitte España. Disponible en línea en: https://www2.deloitte.com/es/es/pages/governance-risk-and-compliance/articles/hacklab-wannaCry.html [Consultado 21-02-2021]

Directiva (UE) 2016/1148 del Parlamento Europeo y del Consejo, de 6 de julio de 2016, Diario Oficial de la Unión Europea, 19/07/2016, 30:194-1. Disponible en línea en: https://www.boe.es/doue/2016/194/L00001-00030.pdf [Consultado 03-05-2021]

Dissmann, Jeffrey & Brechmann, Eike & Czado, Claudia & Kurowicka, Dorota. (2012). Selecting and Estimating Regular Vine Copulas and Application to Financial Returns. Computational Statistics & Data Analysis. 59. 10.1016/j.csda.2012.08.010.

Esteve, M. (2015). *Introducción del Ciber Riesgo en el Mundo Asegurador* (TFM). Disponible en línea en: http://diposit.ub.edu/dspace/bitstream/2445/140345/1/TFM-DEAF-185_Esteve.pdf [Consultado 10-03-2021]

Europol & NATO Strategic Direction South (2018). *Internet Organised Crime Threat Assessment (IOCTA 2018)*. UTB. Disponible en línea en: https://www.europol.europa.eu/sites/default/files/documents/iocta2018.pdf [Consultado 21-02-2021]

Expansión. (2020). *Estados Unidos baja sus tipos de interés*. datosmacro.com. https://datosmacro.expansion.com/tipo-interes/usa [Consultado 14-04-2021]

Gabriela Paoli. (2020). *La hiperconectividad y su influencia en nuestras vidas*. Disponible en línea en: https://www.gabrielapaoli.com/la-hiperconectividad-influencia-nuestras-vidas/#:%7E:text=La%20hiperconectividad%2C%20es%20un%20concepto,internet%20y%20el%20tel%C3%A9fono%20m%C3%B3vil. [Consultado 08-03-2021]

García, G. (2019). *Contratación de la póliza de Ciberriesgos, tratamiento del siniestro y la importancia del reaseguro* (TFM). Disponible en línea en: https://core.ac.uk/download/pdf/237483707.pdf [Consultado 11-03-2021]

Gendre, A. (2020). *Phishers' Favorites: Microsoft reclaims the #1 spot, file phishing spreads, and banks are exploited to harvest email passwords*. VadeSecure. Disponible en línea en: https://www.vadesecure.com/en/blog/phishers-favorites-q1-2020 [Consultado 15-02-2021]

Genest C. y Rivest L.P. (1993). Statistical inference procedures for bivariate Archimedean copulas. J Am Stat Assoc 88:1034–1043

Gordon, L.A., Loeb, M.P. y Sohail, T. (2003). A framework for using insurance for cyber risk management, Communications of the ACM 46, pp. 81-85.

Google y The Cocktail Analysis. (2019). *La ciberseguridad en España. Una perspectiva desde las Pymes, sociedad civil y administración pública*. Disponible en línea en: https://www.ospi.es/export/sites/ospi/documents/documentos/Seguridad-y-privacidad/Google_Panorama-actual-de-la-ciberseguridad-en-Espana.pdf [Consultado 21-02-2021]

Gumbel, E. J. (1960). Bivariate exponential distributions, Journal of the American Statistical Association, 55, 689-707

Deloitte. (2019). *La importancia de contar con un ciberseguro*. Deloitte México. Disponible en línea en: https://www2.deloitte.com/mx/es/pages/dnoticias/articles/ciberseguro.html [Consultado 09-02-2021]

Herath, H.S.B. y Herath, T.C. (2011). Copula-based actuarial model for pricing cyber-insurance policies, Insurance Markets and Companies. Analyses and Actuarial Computations, Vol. 2, No. 1, pp. 7-20

Hofert, M. y Mächler, M. (2011). Nested Archimedean Copulas Meet R: The nacopula Package., Journal of Statistical Software 39(9), 1–20. https://www.jstatsoft.org/v39/i09/

IMF Business School. (2020). *¿Cuáles son los principales organismos relacionados con la ciberseguridad?* Blog de Tecnología - IMF BS. Disponible en línea en: https://blogs.imf-formacion.com/blog/tecnologia/organismos-ciberseguridad-201904/ [Consultado 22-02-2021]

Instituto de Ciencias del Seguro. (2010). *Introducción al Reaseguro*. Fundación Mapfre.

Instituto Nacional de Ciberseguridad. (2017). *Balance Seguridad 2017*. Disponible en línea en: https://www.incibe.es/sites/default/files/paginas/que-hacemos/balance_2017_final_esp.pdf [Consultado 22-02-2021]

Instituto Nacional de Ciberseguridad. (2018). *Balance Seguridad 2018*. Disponible en línea en: https://www.incibe.es/sites/default/files/paginas/que-hacemos/balance_ciberseguridad_2018_esp.pdf [Consultado 22-02-2021]

Instituto Nacional de Ciberseguridad. (2019). *Balance Seguridad 2019*. Disponible en línea en: https://www.incibe.es/sites/default/files/paginas/que-hacemos/balance_ciberseguridad_2019_incibe.pdf [Consultado 22-02-2021]

Insureon. (s. f.). *First-party cyber liability insurance*. Disponible en línea en: https://www.insureon.com/insurance-glossary/cyber-liability-first-party#:%7E:text=First%2Dparty%20coverage%20is%20like,causing%20another%20firm's%20cyber%20losses [Consultado 04-05-2021]

ISO Tools Excellence (2017). Norma ISO 31000: El valor de la gestión de riesgos en las organizaciones. Disponible en línea en: https://www.isotools.org/pdfs-pro/ebook-iso-31000-gestion-riesgos-organizaciones.pdf [Consultado 01-03-2020]

Kolmogorov, A. N. (1933). "Sulla determinizione empirica di una legge di distribuzione" Giorn. Ist. Ital. Attuari , 4, pp. 83–91

Leonor, D. (2021). *Los ciberriesgos se posicionan entre las principales amenazas para las empresas en España*. FÜTURE. Disponible en línea en: https://future.inese.es/los-ciberriesgos-se-posicionan-entre-las-principales-amenazas-para-las-empresas-en-espana/ [Consultado 09-02-2021]

Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales, Boletín Oficial del Estado, 06/12/2018, 294:70. Disponible en línea en: https://www.boe.es/boe/dias/2018/12/06/pdfs/BOE-A-2018-16673.pdf [Consultado 22-02-2021]

Li, J. y Zou, X. (2009), Generalization of the Kermack-McKendrick SIR Model to a Patchy Environment for a Disease with Latency, Mathematical Modelling of Natural Phenomena, Vol. 4, No. 2, pp. 92-118

López, S. (2019). *La breve historia de la ciberseguridad*. Sofistic Cybersecurity. Disponible en línea en: https://www.sofistic.com/blog-ciberseguridad/la-breve-historia-de-la-ciberseguridad/ [Consultado 09-02-2021]

Magallón, E. (2021). *El ataque al SEPE paraliza ya 200.000 citas: ¿cómo se está trabajando?* La Vanguardia. Disponible en línea en: https://www.lavanguardia.com/economia/20210310/6266766/sepe-ataque-informatico-virus-hackeo-citas-erte.html [Consultado 10-03-2021]

NetDiligence. (2020). *Cyber Claims Study 2020 Report*. Disponible en línea en: https://netdiligence.com/wp-content/uploads/2020/11/NetD_2020_Claims_Study_1.1.pdf [Consultado 12-04-2021]

Novales, A. (2017). *Cópulas*. Universidad Complutense de Madrid. Disponible en línea en: https://www.ucm.es/data/cont/media/www/pag-41459/Copulas.pdf [Consultado 10-05-2021]

Orden PCI/487/2019, de 26 de abril, por la que se publica la Estrategia Nacional de Ciberseguridad 2019, aprobada por el Consejo de Seguridad Nacional, Boletín Oficial del Estado, 26/04/2019, 103:19. Disponible en línea en: https://www.boe.es/eli/es/o/2019/04/26/pci487/dof/spa/pdf [Consultado 28-05-2021]

Pacheco, F. (2011). Fuga de información: ¿una amenaza pasajera? Disponible en línea en: https://www.welivesecurity.com/wp-content/uploads/2014/01/fuga_de_informacion.pdf [Consultado 22-04-2021]

Pearson, K. (1895). "Notes on regression and inheritance in the case of two parents". Proceedings of the Royal Society of London. 58: 240–242.

Petrelli, F. (2019). *La cyber insurance per gestire il cyber security risk* (TFM). Disponible en línea en: http://dspace.unive.it/bitstream/handle/10579/16913/871307-1235913.pdf?sequence=2 [Consultado 10-03-2021]

Ponemon Institute LLC. (2018). *2018 Cost of Data Breach Study: Impact of Business Continuity Management*. IBM. https://www.ibm.com/downloads/cas/AEJYBPWA#:~:text=day%20cost%20of%20a%20data,477%20companies%20is%20USD5%2C703. [Consultado 09-04-2021]

RAE. (2021). Ortografía de la Lengua Española, Real Academia Española de la Lengua, Madrid.

Real Decreto-ley 12/2018, de 7 de septiembre, de seguridad de las redes y sistemas de información, Boletín Oficial del Estado, 08/09/2018, 218:24. Disponible en línea en: https://www.boe.es/buscar/pdf/2018/BOE-A-2018-12257-consolidado.pdf [Consultado 03-05-2021]

Reglamento (UE) 2019/881 del Parlamento Europeo y del Consejo, de 17 de abril de 2019, Diario Oficial de la Unión Europea, 07/06/2019, 305:151-15. Disponible en línea en: https://www.boe.es/doue/2019/151/L00015-00069.pdf [Consultado 03-05-2021]

Rinke, K. y Petzoldt, T. (2003). "Modelling the Effects of Temperature and Food on Individual Growth and Reproduction of Daphnia and Their Consequences on the Population Level." Limnologica, 33(4), 293–304

Rinne, H. (2008). The Weibull Distribution, Chapman and Hall/CRC, s.l., 1. edition

Rodríguez, J. (2021). *Misión: cazar a los ciberpiratas*. EL PAÍS. Disponible en línea en: https://elpais.com/eps/2021-02-27/asalto-a-la-fortaleza-digital.html [Consultado 02-03-2021]

Servimedia. (2021). *El sector de la ciberseguridad crecerá este año en España un 8,1%*. Expansión. https://www.expansion.com/economia-digital/2021/02/21/6032396d468aebe7788b45f5.html [Consultado 08-05-2021]

Sklar, A. (1959). Functions de repartition a n dimensions et leurs merges. Publ. Inst. Statist. Univ. Paris 8, pp. 229-231

Sobers, R. (2021). *134 Cybersecurity Statistics and Trends for 2021 | Varonis*. Inside Out Security. https://www.varonis.com/blog/cybersecurity-statistics/ [Consultado 14-04-2021]

Soetaert, K., Petzoldt, T. y Setzer, R.W. (2009). *deSolve: General Solvers for Initial Value Problems of Ordinary Differential Equations (ODE), Partial Differential Equations (PDE), Differential Algebraic Equations (DAE), and Delay Differential Equations (DDE)*. R package version 1.7, URL http://CRAN.R-project.org/package=deSolve

Spearman, C. (1904). "The proof and measurement of association between two things". American Journal of Psychology. 15 (1): 72–101. doi:10.2307/1412159. JSTOR 1412159

Symantec (2018). *Formjacking: Major Increase in Attacks on Online Retailers*. Symantec Blogs. https://symantec-enterprise-blogs.security.com/blogs/threat-intelligence/formjacking-attacks-retailers [Consultado 15-02-2021]

Thaware, V. (2020). *COVID-19 Outbreak Prompts Opportunistic Wave of Malicious Email Campaigns*. Symantec Blogs. https://symantec-enterprise-blogs.security.com/blogs/threat-intelligence/covid-19-outbreak-prompts-opportunistic-wave-malicious-email-campaigns [Consultado 22-04-2021]

Thiber (2016). *La transferencia del ciberriesgo en España*. Disponible en línea en: https://www.thiber.org/ciberseguros.pdf [Consultado 10-03-2021]

United States Census Bureau (2008). Bureau of Justice Statistics Special Report. Cybercrime against Businesses 2005. Disponible en línea en https://www.bjs.gov/content/pub/pdf/cb05.pdf [Consultado 09-04-2021]

ValoraData (2018). *La otra Cara del Internet de las Cosas: Ciberseguridad*. Disponible en línea en https://www.valoradata.com/blog/internet-iot-ciberseguridad/ [Consultado 09-02-2021]

Vodafone (2018). *The Vodafone Cyber Ready Barometer 2018*. Disponible en línea en: https://img.en25.com/Web/VodafoneGroupPLC/%7b1dd2abd4-17b9-4e81-9b23-347f2b41f338%7d_Vodafone-Cyber-Ready-Barometer-research-report-2018.pdf [Consultado 21-02-2021]

Von Mises, R. E. (1928). Wahrscheinlichkeit, Statistik und Wahrheit. Julius Springer.

Wikipedia (2021). *Reaper (antivirus)*. *Wikipedia. La enciclopedia libre*, Wikimedia Foundation, San Francisco, CA. Disponible en línea en: https://es.wikipedia.org/wiki/Reaper_(antivirus) [Consultado 09-02-2021]

Yan, J. (2007). Enjoy the Joy of Copulas: With a Package copula, Journal of Statistical Software. Foundation for Open Access Statistics, Vol. 21, No. 4

7. ANEXO – CÓDIGO R

```
### Gráficos iniciales parte teórica

coron = c(8000,18000,39000,85000)
fecha = c(1,2,3,4)

nombres = c("Feb 27 - Mar 4", "Mar 5 - Mar 11", "Mar 12 - Mar 18", "Mar 19 - Mar 26")

par(mar=c(4.1, 4.1, 2.1, 4.1))
## Figura 2.1
plot(x = fecha, y = coron, pch = 18, cex = 2, cex.axis = 0.75, xaxt = 'n', xlab = "Fe
chas", ylab = "Emails", cex.main = 0.75, col = "lightgoldenrod2", cex.lab = 0.75, cex
.axis = 0.75)
axis(1, at = 1:4, labels= nombres , cex.axis=0.75)

# Representación gráfica 3D cópulas

library(scatterplot3d)
library(rgl)
library(copula)

clay = claytonCopula(2, dim = 3)
set.seed(271)
sample = rCopula(1000, clay)

## Figura 3.1
par(mar=c(4.1, 2.1, 2.1, 4.1))
scatterplot3d(sample, color = "lightgoldenrod2", grid = FALSE, xlab = NA, ylab = NA,
zlab = NA)

## Figura 3.2
scatterplot3d(1 - sample, color = "lightgoldenrod2", grid = FALSE, xlab = NA, ylab =
NA, zlab = NA)

gumbel = gumbelCopula(3, dim = 3)
set.seed(1993)
sample = rCopula(1000, gumbel)

## Figura 3.3
scatterplot3d(sample, color = "lightgoldenrod2", grid = FALSE, xlab = NA, ylab = NA,
zlab = NA)

## Figura 3.4
scatterplot3d(1 - sample, color = "lightgoldenrod2", grid = FALSE, xlab = NA, ylab =
NA, zlab = NA)


### 4.1.2. Modelo SIS

library(simecol)

library(deSolve)

total_pc  = 100 # Total de ordenadores

model_sis = odeModel(main = function(t, y, parms)

{
  p       = parms
  dS      = p["k"]*y["I"]-p["b"]*y["S"]*y["I"] # Ecuación (4.1)
  dI      = -p["k"]*y["I"]+p["b"]*y["S"]*y["I"] # Ecuación (4.2)
  list(c(dS, dI)) },
```

```
times     = c(from = 0, to = 20, by = 0.01),
init      = c(S = total_pc-2, I = 2), # Imposición de valores iniciales
parms     = c(k = 1, b = 0.01), # Imposición tasa γ y tasa β
solver    = "lsoda")

simulacion_sis = sim(model_sis)

par(mar=c(6.1, 4.1, 5.1, 1.1))

## Figuras 4.2, 4.3 y 4.4
plot(simulacion_sis, xlab = "Tiempo (horas)", ylab = "Ordenadores", col.main = "light
goldenrod3", text.cell=list(cex=0.5), cex.lab = 0.75, cex.axis = 0.75)


### 4.1.3. Modelo SIR

model_sir = function(time, variables, parameters) {

  with(as.list(c(variables, parameters)), {
    dS = -beta * I * S # Ecuación (4.5)
    dI =  beta * I * S - gamma * I # Ecuación (4.6)
    dR =  gamma * I # Ecuación (4.7)
    return(list(c(dS, dI, dR))) }) }

param   = c(
  beta  = 0.04, # Tasa de infección de cada ordenador por hora
  gamma = 0.5) # Tasa de recuperación por hora

inicial_valores = c(
  S = 99, # Estado inicial Susceptibles
  I = 1, # Estado inicial Infectados
  R = 0) # Estado inicial Recuperados

tiempo_horas = seq(0, 12) # Horas

valor_model_sir = ode(y = inicial_valores, times = tiempo_horas, func = model_sir, pa
rms = param)

valor_model_sir = as.data.frame(valor_model_sir)

## Figura 4.6
par(mar=c(4.1, 4.1, 2.1, 4.1))

with(valor_model_sir, plot(tiempo_horas, S, type = "l", col = "lightgoldenrod2", xlab
= "Tiempo (horas)", ylab = "Número de ordenadores", lwd = 2, cex.lab = 0.75, cex.axis
= 0.75))
with(valor_model_sir,lines(tiempo_horas, I, col = "aquamarine2", type ="l", lwd = 2))
with(valor_model_sir, lines(tiempo_horas, R, col = "plum3", type ="l", lwd = 2))
legend("right", c("Susceptibles", "Infectados", "Recuperados"),
       col = c("lightgoldenrod2", "aquamarine2", "plum3"), lty = 1, bty = "n", pch =
16, cex = 0.65)


### 4.2. Funciones Cópula


library(MASS)
library(plot.matrix)
library(viridis)

a     = 3
b     = 2000
matriz_sig = matrix(c(1, 0.2, 0.1,
                      0.2, 1, -0.4,
                      0.1, -0.4, 1),
                    nrow=3)
```

```r
z = mvrnorm(b, mu = rep(0, a), Sigma= matriz_sig, empirical = T)

z2 = z[1:7,1:3]

set.seed(93)
par(mar=c(2.1, 2.1, 1.1, 4.1))
## Figura 4.7
plot(z2, cex.main = 0.75, main = NA, xlab = "Columnas", ylab = "Filas", col = viridis
(6),  digits=2, text.cell=list(cex=0.5), cex.lab = 0.75, cex.axis = 0.75)

round(cor(z, method='spearman'), digits = 2)

library(psych)
## Figura 4.8
pairs.panels(z, hist.col = "lightskyblue3", method = "spearman", density = TRUE, elli
pses = TRUE, col = "aliceblue", pch = 21, bg=c("lightgoldenrod2", "aquamarine2", "plu
m3"), cex.cor = 0.5, labels = c("Var 1", "Var 2", "Var 3"), cex.labels = 1.3)

u = pnorm(z)

## Figura 4.9
pairs.panels(u, hist.col = "lightskyblue3", method = "spearman", density = TRUE, elli
pses = TRUE, col = "aliceblue", pch = 21, bg=c("lightgoldenrod2", "aquamarine2", "plu
m3"), cex.cor = 0.5, labels = c("Var 1", "Var 2", "Var 3"), cex.labels = 1.3)

unif_1 = u[,1]
unif_2 = u[,2]
unif_3 = u[,3]

## Figura 4.10
plot3d(unif_1,unif_2,unif_3, pch = 20, col='lightskyblue3')

Gamma = qgamma(u[,1],shape=2,scale=1)
Beta = qbeta(u[,2],2,2)
tStudent = qt(u[,3],df=5)

muestra = cbind(Gamma, Beta, tStudent)
## Figura 4.11
pairs.panels(muestra, hist.col = "lightskyblue3", method = "spearman", density = TRUE
, ellipses = TRUE, col = "aliceblue", pch = 21, bg=c("lightgoldenrod2", "aquamarine2"
, "plum3"), cex.cor = 0.5, labels = c("Var 1", "Var 2", "Var 3"), cex.labels = 1.3)


library(rgl)
## Figura 4.12
plot3d(Gamma,Beta,tStudent, pch = 20, col='lightskyblue3')

# Mismo procedimiento pero con cópulas

set.seed(100)
copula_normal = normalCopula(param=c(0.2,0.1,-0.4), dim = 3, dispstr = "un")
distr_mult    = mvdc(copula=copula_normal, margins=c("gamma", "beta", "t"),  paramMar
gins=list(list(shape=2, scale=1), list(shape1=2, shape2=2), list(df=5)))
muestra_cop       = rMvdc(1000, distr_mult)
colnames(muestra_cop) = c("a", "b", "c")
## Figura 4.13
pairs.panels(muestra_cop, hist.col = "lightskyblue3", method = "spearman", density =
TRUE, ellipses = TRUE, col = "aliceblue", pch = 21, bg=c("lightgoldenrod2", "aquamari
ne2", "plum3"), cex.cor = 0.5, labels = c("Var 1", "Var 2", "Var 3"), cex.labels = 1.
3)


### 4.3.3. Tarificación del ciberseguro con datos de 2003

# En primer lugar, definimos las dos variables de nuestro estudio, los ordenadores af
```

```
ectados y las pérdidas con datos de ICSA (Bridwell, 2004; en Hearth y Herath, 2011)
## Tabla 4.1
ord = c(1291, 849, 238, 140, 118, 108, 87, 70, 63, 58, 50, 47, 21, 21, 19)
perd = c(355648.72, 339832.66, 115729.51, 65090.38, 45402.25, 66053.73, 39182.88, 195
56.82, 20087.13, 20465.35, 10180.13, 11769.29, 6944.48, 5339.08, 7547.77)

par(mar=c(4.1, 4.1, 2.1, 4.1))
## Figura 4.14
plot(x = ord, y = perd, pch = 18, cex = 2, xlab = "Ordenadores", ylab = "Pérdidas", c
ex.main = 0.75, col = "lightgoldenrod2", text.cell=list(cex=0.5), cex.lab = 0.75, cex
.axis = 0.75)

library(VineCopula)

library(copula)

# Encontramos la cópula que mejor se adapta a nuestros datos
## Ecuación (4.15)
u = pobs(as.matrix(cbind(ord,perd)))[,1]
v = pobs(as.matrix(cbind(ord,perd)))[,2]
cop_2003 = BiCopSelect(u,v, familyset=NA)
cop_2003

## Ecuación (4.16)
tau(claytonCopula(param = 10)) # Para comprobar que la tau está bien estimada

# Ajuste Distribuciones sobre variable Ordenadores

library(fitdistrplus)
library(actuar)

ajuste_ord_weibull = fitdist(ord, "weibull")
ajuste_ord_pareto = fitdist(ord, "pareto", start = list(shape = 1, scale = 500))
ajuste_ord_exp = fitdist(ord, "exp")
ajuste_ord_gamma = fitdist(ord, "gamma")

## Figura 4.15
par(mfrow=c(2,2), mar=c(2.1, 2, 0.1, 0.5))
plot.legend = c("Weibull", "Pareto", "Exponencial", "Gamma")
denscomp(list(ajuste_ord_weibull, ajuste_ord_pareto, ajuste_ord_exp, ajuste_ord_gamma
), legendtext = plot.legend, ylab = NA, cex = 0.7, main = NA)
plot.legend = c("Weibull", "Pareto", "Exponencial", "Gamma")
qqcomp(list(ajuste_ord_weibull, ajuste_ord_pareto, ajuste_ord_exp, ajuste_ord_gamma),
legendtext = plot.legend, ylab = NA, cex = 0.7, main = NA)
plot.legend = c("Weibull", "Pareto", "Exponencial", "Gamma")
cdfcomp(list(ajuste_ord_weibull, ajuste_ord_pareto, ajuste_ord_exp, ajuste_ord_gamma)
, legendtext = plot.legend, ylab = NA, cex = 0.7, main = NA)
plot.legend = c("Weibull", "Pareto", "Exponencial", "Gamma")
ppcomp(list(ajuste_ord_weibull, ajuste_ord_pareto, ajuste_ord_exp, ajuste_ord_gamma),
legendtext = plot.legend, ylab = NA, cex = 0.7, main = NA)

# Estadísticos para ver qué distribución es la que mejor se ajusta
## Tablas 4.2 y 4.3
gofstat(list(ajuste_ord_weibull, ajuste_ord_pareto, ajuste_ord_exp, ajuste_ord_gamma)
, fitnames = c("Weibull", "Pareto", "Exponencial", "Gamma"))

# Parece ser que la Pareto es la que mejor se ajusta

# Ajuste Distribuciones sobre variable Pérdidas

ajuste_perd_weibull = fitdist(perd, "weibull")
ajuste_perd_pareto = fitdist(perd, "pareto", start = list(shape = 1, scale = 500))
#ajuste_perd_exp = fitdist(perd/1000, "exp")
#ajuste_perd_gamma = fitdist(perd/1000, "gamma")

# Debemos dividir la exponencial y la gamma entre 1.000, pues al ser números grandes
```

```
el programa los eleva a la exponencial y se van al infinito
## Figura 4.16
par(mfrow=c(2,2), mar=c(2.1, 2, 0.1, 0.5))
plot.legend = c("Weibull", "Pareto")
denscomp(list(ajuste_perd_weibull, ajuste_perd_pareto), legendtext = plot.legend, yla
b = NA, cex = 0.7, main = NA)
plot.legend = c("Weibull", "Pareto")
qqcomp(list(ajuste_perd_weibull, ajuste_perd_pareto), legendtext = plot.legend, ylab
= NA, cex = 0.7, main = NA)
plot.legend = c("Weibull", "Pareto")
cdfcomp(list(ajuste_perd_weibull, ajuste_perd_pareto), legendtext = plot.legend, ylab
= NA, cex = 0.7, main = NA)
plot.legend = c("Weibull", "Pareto")
ppcomp(list(ajuste_perd_weibull, ajuste_perd_pareto), legendtext = plot.legend, ylab
= NA, cex = 0.7, main = NA)

# Estadísticos para ver qué distribución es la que mejor se ajusta
## Tablas 4.4 y 4.5
gofstat(list(ajuste_perd_pareto, ajuste_perd_weibull), fitnames = c("Weibull", "Paret
o"))


# Construimos las distribuciones multivariantes con la estructura de dependencia que
seguirán estas a través de la cópula mencionada

distr_cop_pareto = mvdc(copula= claytonCopula(10), margins = c("pareto","pareto"),
                        paramMargins=list(list(shape = 1.701058 , scale = 164.689324)
,
                                          list(shape = 1.621569 , scale = 57035.47220
4)))

set.seed(349)
muestra_cop_pareto = rMvdc(300, distr_cop_pareto) # Crea una muestra bivariante de nu
estras dos variables

# Comparamos los valores observados con los simulados
## Figura 4.17
par(mar=c(4.1, 4.1, 2.1, 4.1))
plot(x = ord, y = perd, pch = 18, cex = 2, xlab = "Ordenadores", ylab = "Pérdidas", c
ex.main = 0.75, col = "lightgoldenrod2", text.cell=list(cex=0.5), cex.lab = 0.75, cex
.axis = 0.75)
points(muestra_cop_pareto[,1],muestra_cop_pareto[,2], col='lightskyblue3', pch=21)
legend('bottomright',c('Valor Observado','Valor Simulado'), col=c('lightgoldenrod2','
lightskyblue3'), lty = 1, bty = "n", pch = 16, cex = 0.65)

# Comparación correlaciones entre muestra observada y simulada
## Tabla 4.7
round(cor(ord, perd, method='spearman'), d = 4) # Observada
round(cor(muestra_cop_pareto[,1], muestra_cop_pareto[,2], method='spearman'), d = 4)
# Simulada

## Ecuación (4.10) para obtener la distribución de pérdidas de la suma asegurada:

distr_perdidas = seq(1,300)

l = min(ord)
m = max(ord)
a1 = 400
a2 = 125
a3 = 300

for(i in 1:300){

  if(muestra_cop_pareto[i,1]<l) {distr_perdidas[i] = a1}
  else if(muestra_cop_pareto[i,1] >= l && muestra_cop_pareto[i,1]<m) {distr_perdidas[
i] = a2 + ((muestra_cop_pareto[i,1]-l)/muestra_cop_pareto[i,1])*(muestra_cop_pareto[i
```

```
,2]/10)}
  else {distr_perdidas[i] = a3 +((muestra_cop_pareto[i,1]-m)/muestra_cop_pareto[i,1])
*(muestra_cop_pareto[i,2]/10)}

}

## Figura 4.18
hist(distr_perdidas, xlab = "Pérdidas", ylab = "Densidad", col = "lightgoldenrod2", f
req = FALSE, main = NA, cex.axis = 0.75, cex.lab = 0.75)
#lines(density(distr_perdidas), col ="black", lwd = 1)

# a)    Póliza básica first party damage sin franquicia. Ecuación (4.12)

prima_2003_a = round(mean(0.052 * distr_perdidas * exp(x=-197/365*0.0125)), d = 2)

# Bucle para hallar la prima media y la suma asegurada media

dim = 10000

prima_2003_a_bucle = seq(1,dim)
suma_aseg_2003 = seq(1,dim)


for(j in 1:dim){

  muestra_cop_pareto_bucle = rMvdc(300, distr_cop_pareto)

  distr_perdidas_bucle = seq(1,300)

  for(i in 1:300){

    if(muestra_cop_pareto_bucle[i,1]<l) {distr_perdidas_bucle[i] = a1}
    else if(muestra_cop_pareto_bucle[i,1] >= l && muestra_cop_pareto_bucle[i,1]<m) {d
istr_perdidas_bucle[i] = a2 + ((muestra_cop_pareto_bucle[i,1]-l)/muestra_cop_pareto_b
ucle[i,1])*(muestra_cop_pareto_bucle[i,2]/10)}
    else {distr_perdidas_bucle[i] = a3 +((muestra_cop_pareto_bucle[i,1]-m)/muestra_co
p_pareto_bucle[i,1])*(muestra_cop_pareto_bucle[i,2]/10)}

  }


  suma_aseg_2003[j] = mean(distr_perdidas_bucle)
  prima_2003_a_bucle[j] = mean(0.052 * distr_perdidas_bucle * exp(x=-197/365*0.0125))

}

mean(prima_2003_a_bucle) # Esta será la prima promedio
mean(suma_aseg_2003) # Esta será la suma asegurada promedio


# Bucle para saber qué semilla fijar para que salga la prima promedio estimada

l = min(ord)
m = max(ord)
a1 = 400
a2 = 125
a3 = 300


for(k in seq(300,350,1)){

  print(k)
  set.seed(k)

  muestra_cop_pareto_bucle2 = rMvdc(300, distr_cop_pareto)
  distr_perdidas_bucle2 = seq(1,300)
```

```
  for(i in 1:300){

    if(muestra_cop_pareto_bucle2[i,1]<l) {distr_perdidas_bucle2[i] = a1}
    else if(muestra_cop_pareto_bucle2[i,1] >= l && muestra_cop_pareto_bucle2[i,1]<m)
{distr_perdidas_bucle2[i] = a2 + ((muestra_cop_pareto_bucle2[i,1]-l)/muestra_cop_pare
to_bucle2[i,1])*(muestra_cop_pareto_bucle2[i,2]/10)}
    else {distr_perdidas_bucle2[i] = a3 +((muestra_cop_pareto_bucle2[i,1]-m)/muestra_
cop_pareto_bucle2[i,1])*(muestra_cop_pareto_bucle2[i,2]/10)}

  }

  prima_2003_a_bucle2 = mean(0.052 * distr_perdidas_bucle2 * exp(x=-197/365*0.0125))

  print(prima_2003_a_bucle2)

  if(prima_2003_a_bucle2 >= 400 && prima_2003_a_bucle2 < 407) {break}

}

# b)    Póliza first party damage con franquicia. Primero aplicamos la Ecuación (4.13
) para obtener la distribución de pérdidas:

distr_perdidas_b = seq(1,300)
prima_2003_b = seq(1,length(seq(0,4000,500)))
j = 0

for(franq in seq(0,4000,500)){

  j = j + 1
  d = franq

  for(i in 1:300){

    if(distr_perdidas[i]<=d) {distr_perdidas_b[i] = 0}
    else {distr_perdidas_b[i] = distr_perdidas[i] - d}

  }

  prima_2003_b[j] = mean(0.052 * distr_perdidas_b * exp(x=-197/365*0.0125))

}

## Tabla 4.8
round(prima_2003_b, d = 2)

library("writexl")
write_xlsx(as.data.frame(round(prima_2003_b, d = 2)),"C:\\Users\\ayaqueagui001\\Deskt
op\\TFM\\prima_b_2003.xlsx")

# c)    Póliza first party damage con coaseguro, franquicia y límite. Ecuación (4.14)

distr_perdidas_c = seq(1,300)
prima_2003_c = seq(1,length(seq(0,2500,500)))
matriz_prima_c = matrix(nrow = 4, ncol = 6)
k = 20000 # límite
m = 0

for(coas in seq(0.05,0.2,0.05)){

  a = coas
  m = m + 1 # indicador fila
  j = 0

  for(franq in seq(0,2500,500)){
```

```
    j = j + 1 # posición vector
    d = franq

    for(i in 1:300){

      if(distr_perdidas[i]<=d) {distr_perdidas_c[i] = 0}
      else if(distr_perdidas[i]<d+k/(1-a)) {distr_perdidas_c[i]= (1-a)*(distr_perdida
s[i]-d)}
      else {distr_perdidas_c[i] = k}

    }

    prima_2003_c[j] = mean(0.052 * distr_perdidas_c * exp(x=-197/365*0.0125))
    matriz_prima_c[m,j] = prima_2003_c[j]

  }

}

## Tabla 4.9
round(matriz_prima_c, d = 2)

write_xlsx(as.data.frame(matriz_prima_c),"C:\\Users\\ayaqueagui001\\Desktop\\TFM\\pri
ma_c10000_2003.xlsx")

### 4.3.4. Tarificación del ciberseguro con datos de 2020

# En primer lugar, definimos las dos variables de nuestro estudio, los ordenadores af
ectados y las pérdidas con datos de NetDiligence (2020)
## Tabla 4.10
ord2 = c(318, 430, 606, 56, 127, 100, 176, 5, 31, 258, 20, 915, 161, 782, 157, 344, 1
2, 10, 15, 9, 191, 2)
perd2 = c(41.9, 272.7, 299, 7.2, 6.4, 5, 26.3, 0.389, 1.5, 17.8, 7.1, 130.5, 12.3, 79
.4, 11.2, 5.7, 19.4, 0.381, 0.4, 1.3, 26.8, 0.091)*1000

par(mar=c(4.1, 4.1, 2.1, 4.1))
## Figura 4.19
plot(x = ord2, y = perd2, pch = 18, cex = 2, xlab = "Ordenadores", ylab = "Pérdidas",
cex.main = 0.75, col = "lightgoldenrod2", text.cell=list(cex=0.5), cex.lab = 0.75, ce
x.axis = 0.75)

library(VineCopula)

# Encontramos la cópula que mejor se adapta a nuestros datos
## Ecuación (4.17)
u2 = pobs(as.matrix(cbind(ord2,perd2)))[,1]
v2 = pobs(as.matrix(cbind(ord2,perd2)))[,2]
cop_2020 = BiCopSelect(u2,v2, familyset=NA)
cop_2020

## Ecuación (4.16)

library(copula)
tau(gumbelCopula(param = 3.38)) # Para comprobar que la tau está bien estimada

# Ajuste Distribuciones sobre variable Ordenadores

library(fitdistrplus)
library(actuar)

ajuste_ord_weibull2 = fitdist(ord2, "weibull")
ajuste_ord_pareto2 = fitdist(ord2, "pareto", start = list(shape = 1, scale = 500))
ajuste_ord_exp2 = fitdist(ord2, "exp")
ajuste_ord_gamma2 = fitdist(ord2, "gamma")
```

```
par(mfrow=c(2,2), mar=c(2.1, 2, 0.1, 0.5))
## Figura 4.20
plot.legend = c("Weibull", "Pareto", "Exponencial", "Gamma")
denscomp(list(ajuste_ord_weibull2, ajuste_ord_pareto2, ajuste_ord_exp2, ajuste_ord_ga
mma2), legendtext = plot.legend, ylab = NA, cex = 0.7, main = NA)
plot.legend = c("Weibull", "Pareto", "Exponencial", "Gamma")
qqcomp(list(ajuste_ord_weibull2, ajuste_ord_pareto2, ajuste_ord_exp2, ajuste_ord_gamm
a2), legendtext = plot.legend, ylab = NA, cex = 0.7, main = NA)
plot.legend = c("Weibull", "Pareto", "Exponencial", "Gamma")
cdfcomp(list(ajuste_ord_weibull2, ajuste_ord_pareto2, ajuste_ord_exp2, ajuste_ord_gam
ma2), legendtext = plot.legend, ylab = NA, cex = 0.7, main = NA)
plot.legend = c("Weibull", "Pareto", "Exponencial", "Gamma")
ppcomp(list(ajuste_ord_weibull2, ajuste_ord_pareto2, ajuste_ord_exp2, ajuste_ord_gamm
a2), legendtext = plot.legend, ylab = NA, cex = 0.7, main = NA)

# Estadísticos para ver qué distribución es la que mejor se ajusta
## Tablas 4.11 y 4.12
gofstat(list(ajuste_ord_weibull2, ajuste_ord_pareto2, ajuste_ord_exp2, ajuste_ord_gam
ma2), fitnames = c("Weibull", "Pareto", "Exponencial", "Gamma"))

# Parece ser que la Gamma es la que mejor se ajusta

# Ajuste Distribuciones sobre variable Pérdidas

ajuste_perd_weibull2 = fitdist(perd2, "weibull")
ajuste_perd_pareto2 = fitdist(perd2, "pareto", start = list(shape = 1, scale = 500))
#ajuste_perd_exp = fitdist(perd/1000, "exp")
#ajuste_perd_gamma = fitdist(perd/1000, "gamma")

# Debemos dividir la exponencial y la gamma entre 1.000 pues al ser números grandes e
l programa los eleva a la exponencial y se van al infinito

## Figura 4.21
par(mfrow=c(2,2), mar=c(2.1, 2, 0.1, 0.5))
plot.legend = c("Weibull", "Pareto")
denscomp(list(ajuste_perd_weibull2, ajuste_perd_pareto2), legendtext = plot.legend, y
lab = NA, cex = 0.7, main = NA)
plot.legend = c("Weibull", "Pareto")
qqcomp(list(ajuste_perd_weibull2, ajuste_perd_pareto2), legendtext = plot.legend, yla
b = NA, cex = 0.7, main = NA)
plot.legend = c("Weibull", "Pareto")
cdfcomp(list(ajuste_perd_weibull2, ajuste_perd_pareto2), legendtext = plot.legend, yl
ab = NA, cex = 0.7, main = NA)
plot.legend = c("Weibull", "Pareto")
ppcomp(list(ajuste_perd_weibull2, ajuste_perd_pareto2), legendtext = plot.legend, yla
b = NA, cex = 0.7, main = NA)

# Estadísticos para ver qué distribución es la que mejor se ajusta
## Tablas 4.13 y 4.14
gofstat(list(ajuste_perd_weibull2, ajuste_perd_pareto2), fitnames = c("Weibull", "Par
eto"))

# Construimos las distribuciones multivariantes con la estructura de dependencia que
seguirán estas a través de la cópula mencionada

distr_cop_gumbel = mvdc(copula= gumbelCopula(3.38), margins = c("gamma","weibull"),
                        paramMargins=list(list(shape = 0.602019645, rate = 0.00280246
3),
                                          list(shape = 5.252395e-01, scale = 2.304491
e+04)))
set.seed(638)
muestra_cop_gumbel = rMvdc(300, distr_cop_gumbel)

# Comparamos los valores observados con los simulados

par(mar=c(4.1, 4.1, 2.1, 4.1))
```

```r
## Figura 4.22
plot(x = ord2, y = perd2, pch = 18, cex = 2, xlab = "Ordenadores", ylab = "Pérdidas",
cex.main = 0.75, col = "lightgoldenrod2", text.cell=list(cex=0.5), cex.lab = 0.75, ce
x.axis = 0.75)
points(muestra_cop_gumbel[,1],muestra_cop_gumbel[,2], col='lightskyblue3', pch=21)
legend('bottomright',c('Valor Observado','Valor Simulado'), col=c('lightgoldenrod2','
lightskyblue3'), lty = 1, bty = "n", pch = 16, cex = 0.65)

# Comparación correlaciones entre la muestra observada y simulada
## Tabla 4.16
round(cor(ord2, perd2, method='spearman'), d = 4) # Muestra observada
round(cor(muestra_cop_gumbel[,1], muestra_cop_gumbel[,2], method='spearman'), d = 4)
# MUestra simulada

## Ecuación (4.10) para obtener la distribución de pérdidas:

distr_perdidas2 = seq(1,300)

l = min(ord2)
m = max(ord2)
a1 = 400
a2 = 125
a3 = 300

for(i in 1:300){

  if(muestra_cop_gumbel[i,1]<l) {distr_perdidas2[i] = a1}
  else if(muestra_cop_gumbel[i,1] >= l && muestra_cop_gumbel[i,1]<m) {distr_perdidas2
[i] = a2 + ((muestra_cop_gumbel[i,1]-l)/muestra_cop_gumbel[i,1])*(muestra_cop_gumbel[
i,2]/10)}
  else {distr_perdidas2[i] = a3 +((muestra_cop_gumbel[i,1]-m)/muestra_cop_gumbel[i,1]
)*(muestra_cop_gumbel[i,2]/10)}

}

## Figura 4.23
hist(distr_perdidas2, xlab = "Pérdidas", ylab = "Densidad", col = "lightgoldenrod2",
freq = FALSE, main = NA, cex.lab = 0.75, cex.axis = 0.75)
#lines(density(distr_perdidas2), col ="black", lwd = 1)

# a)    Póliza básica first party damage sin franquicia. Ecuación (4.12)

prima_2020_a = round(mean(0.25 * distr_perdidas2 * exp(x=-207/365*0)), d = 2)

round(((prima_2020_a/prima_2003_a)-1)*100, d = 2) # Aumento con respecto a la prima d
e 2003

# Bucle para hallar la prima media

dim = 10000

prima_2020_a_bucle = seq(1,dim)
suma_aseg_2020 = seq(1,dim)


for(j in 1:dim){

  muestra_cop_gumbel_bucle = rMvdc(300, distr_cop_gumbel)

  distr_perdidas_bucle = seq(1,300)

  for(i in 1:300){

    if(muestra_cop_gumbel_bucle[i,1]<l) {distr_perdidas_bucle[i] = a1}
    else if(muestra_cop_gumbel_bucle[i,1] >= l && muestra_cop_gumbel_bucle[i,1]<m) {d
istr_perdidas_bucle[i] = a2 + ((muestra_cop_gumbel_bucle[i,1]-l)/muestra_cop_gumbel_b
```

```
ucle[i,1])*(muestra_cop_gumbel_bucle[i,2]/10)}
    else {distr_perdidas_bucle[i] = a3 +((muestra_cop_gumbel_bucle[i,1]-m)/muestra_co
p_gumbel_bucle[i,1])*(muestra_cop_gumbel_bucle[i,2]/10)}

  }


  suma_aseg_2020[j] = mean(distr_perdidas_bucle)
  prima_2020_a_bucle[j] = mean(0.25 * distr_perdidas_bucle * exp(x=-207/365*0))

}


mean(suma_aseg_2020)
mean(prima_2020_a_bucle)
mean(((prima_2020_a_bucle/prima_2003_a_bucle)-1)*100)

# Bucle para saber qué semilla fijar para que salga la prima promedio estimada

l = min(ord2)
m = max(ord2)
a1 = 400
a2 = 125
a3 = 300


for(k in seq(569,100000,1)){

  print(k)
  set.seed(k)

  muestra_cop_gumbel_bucle2 = rMvdc(300, distr_cop_gumbel)
  distr_perdidas_bucle2 = seq(1,300)


  for(i in 1:300){

    if(muestra_cop_gumbel_bucle2[i,1]<l) {distr_perdidas_bucle2[i] = a1}
    else if(muestra_cop_gumbel_bucle2[i,1] >= l && muestra_cop_gumbel_bucle2[i,1]<m)
{distr_perdidas_bucle2[i] = a2 + ((muestra_cop_gumbel_bucle2[i,1]-l)/muestra_cop_gumb
el_bucle2[i,1])*(muestra_cop_gumbel_bucle2[i,2]/10)}
    else {distr_perdidas_bucle2[i] = a3 +((muestra_cop_gumbel_bucle2[i,1]-m)/muestra_
cop_gumbel_bucle2[i,1])*(muestra_cop_gumbel_bucle2[i,2]/10)}

  }

  prima_2020_a_bucle2 = mean(0.25 * distr_perdidas_bucle2 * exp(x=-207/365*0))

  print(prima_2020_a_bucle2)

  if(prima_2020_a_bucle2 >= 864 && prima_2020_a_bucle2 < 867) {break}

}

# b)    Póliza first party damage con franquicia. Primero aplicamos la Ecuación (4.13
) para obtener la distribución de pérdidas:

distr_perdidas_b = seq(1,300)
prima_2020_b = seq(1,length(seq(0,4000,500)))
j = 0

for(franq in seq(0,4000,500)){

  j = j + 1
  d = franq
```

```
  for(i in 1:300){

    if(distr_perdidas2[i]<=d) {distr_perdidas_b[i] = 0}
    else {distr_perdidas_b[i] = distr_perdidas2[i] - d}

  }

  prima_2020_b[j] = mean(0.25 * distr_perdidas_b * exp(x=-207/365*0))

}

## Tabla 4.17
round(prima_2020_b, d = 2)
round(mean(((prima_2020_b/prima_2003_b)-1)*100), d = 2) # Subida media

write_xlsx(as.data.frame(prima_2020_b),"C:\\Users\\ayaqueagui001\\Desktop\\TFM\\prima
_b_2020.xlsx")


# c)    Póliza first party damage con coaseguro, franquicia y límite. Ecuación (4.14)

distr_perdidas_c = seq(1,300)
prima_2020_c = seq(1,length(seq(0,2500,500)))
matriz_prima_c = matrix(nrow = 4, ncol = 6)
k = 20000 # límite
m = 0

for(coas in seq(0.05,0.2,0.05)){

  a = coas
  m = m + 1 # indicador fila
  j = 0

  for(franq in seq(0,2500,500)){

    j = j + 1 # posición vector
    d = franq

    for(i in 1:300){

      if(distr_perdidas2[i]<=d) {distr_perdidas_c[i] = 0}
      else if(distr_perdidas2[i]<d+k/(1-a)) {distr_perdidas_c[i]= (1-a)*(distr_perdid
as2[i]-d)}
      else {distr_perdidas_c[i] = k}

    }

    prima_2020_c[j] = mean(0.25 * distr_perdidas_c * exp(x=-207/365*0))
    matriz_prima_c[m,j] = prima_2020_c[j]

  }

}

## Tabla 4.18
round(matriz_prima_c, d = 2)
write_xlsx(as.data.frame(matriz_prima_c),"C:\\Users\\ayaqueagui001\\Desktop\\TFM\\pri
ma_c10000_2020.xlsx")
```