LUISS

Dipartimento
di Economia e Finanza

Cattedra di Computational Tools for Finance

Modelli quantitativi per la valutazione del Cyber Risk

Prof. Valerio Marchisio
RELATORE

Prof. Sara Biagini
CORRELATORE

Biagio Miele 724891
CANDIDATO

Anno Accademico 2020/2021

A mia madre, a mio padre
al loro infinito amore
per avermi insegnato a pensare
liberamente

A mio fratello
perché solo il sangue che ci lega
posso chiamarlo casa

Sommario

Introduzione

La progressiva digitalizzazione dell'economia globale a cui abbiamo assistito a partire dalla metà del secolo scorso ha posto nuove sfide e creato nuove opportunità per le imprese di qualsiasi settore. Ai consueti beni capitali "fisici", posti al *core* del modello economico tradizionale, si è progressivamente affiancata una nuova tipologia di capitale: i beni "intangibili", sottoforma di dati, informazioni e tecnologie. Questo processo è stato ovviamente incoraggiato e agevolato dallo sviluppo del settore delle telecomunicazioni e in particolare dell'informatica, che ha permesso agli agenti economici di conservare, scambiare, elaborare e analizzare dati e informazioni in quantità sempre più massicce e modalità sempre più raffinate.

Se da un lato il crescente grado di informatizzazione e digitalizzazione delle imprese fornisce ad esse nuove potenzialità, dall'altro le espone tuttavia a una nuova tipologia di rischio, definita *cyber risk*, connessa principalmente alla perdita della disponibilità o alla compromissione dell'integrità dei dati e/o dei sistemi di elaborazione delle informazioni e alle conseguenze di tali eventi. Come vedremo, per il particolare ruolo ricoperto all'interno del sistema economico e per la natura dell'attività svolta, il settore finanziario risulta tra i più esposti al *cyber risk*.

Si tratta dunque di una categoria di rischio che sta velocemente attirando attenzione sia dal mondo accademico che dal mondo dell'industria. L'annuale "Future Risks Report" pubblicato da AXA, nel 2021, cita il *cyber risk* al secondo posto della Top 10 dei rischi emergenti per il prossimo decennio, dietro solo al rischio connesso al cambiamento climatico e, sorprendentemente, davanti al rischio portato da pandemie e malattie infettive (Figura 1).

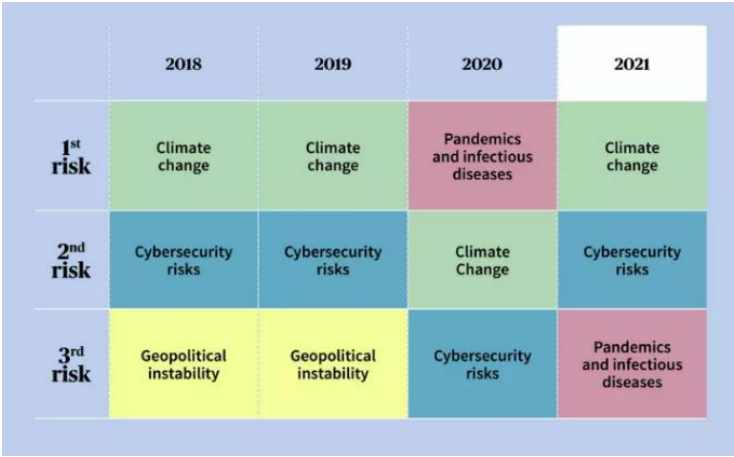

Figura 1. La top 3 dei rischi dal 2018 al 2021 secondo il sondaggio annuale di AXA. Fonte: AXA's Future Risks Report 2021.

Nonostante la forte espansione del fenomeno, che ha conosciuto un ulteriore incremento nei mesi più drammatici della pandemia Covid-19, la ricerca a riguardo è ancora limitata,

ostacolata sia da una significativa mancanza di dataset utilizzabili che da intrinseche difficoltà di modellazione del problema che saranno più avanti esaminate.

Il presente elaborato punta, dunque, a fornire al lettore un'analisi estensiva del fenomeno, passando dal generale al particolare con un focus sul settore finanziario e proponendo un'applicazione pratica di tecniche di quantificazione e gestione di questa forma di rischio in continua evoluzione. Nel primo Capitolo verrà introdotto e illustrato il fenomeno, analizzandone le caratteristiche intrinseche, le modalità in cui il *cyber risk* si materializza, quali comparti dell'economia sono più esposti e le risposte del legislatore europeo all'evoluzione del problema. Nel secondo Capitolo, invece, è presentata una panoramica delle metodologie implementate nel sistema di Basilea II e III per la valutazione e gestione del rischio operativo, di cui, come vedremo, il *cyber risk* rappresenta un sottoinsieme; in particolare, verrà illustrato il *Loss Distribution Approach*, una metodologia derivata dalle scienze attuariali che, al momento, rappresenta il *gold standard* per queste categorie di rischi e la loro quantificazione. Infine, nel terzo Capitolo è proposta un'applicazione della metodologia, volta alla creazione di un modello quantitativo che possa catturare al meglio le caratteristiche del *cyber risk* e restituire delle misure di rischio, *VaR* ed *Expected Shortfall*, rappresentative del profilo di rischio *cyber* del settore finanziario (e di altri settori, analizzati nel paragrafo 3.4.1).

Capitolo 1

Le caratteristiche del *Cyber Risk*

1.1 Definizione

Non è compito agevole individuare una definizione di *cyber risk* universale, in quanto la ricerca scientifica a riguardo non ha ancora adottato uno standard[1]. Si tratta in effetti di una categoria di rischio complessa, che per essere trattata adeguatamente presuppone conoscenze sia tecniche che economiche e che presenta un forte grado di interdisciplinarità: contributi sull'argomento arrivano, tra gli altri, dai campi della finanza, delle assicurazioni, del risk management, ma anche dall'informatica e ovviamente dalla ricerca sulla cybersecurity.

Cebula e Young (2010), ad esempio, definiscono i *cyber security risks* come «rischi operativi agli asset informativi e tecnologici con conseguenze riguardanti la confidenzialità, disponibilità e integrità delle informazioni o dei sistemi informatici», definizione ripresa anche da Bouveret (2018). Similarmente, Eling e Schnell (2016) definiscono *cyber risk* «ogni rischio emergente dall'uso di tecnologie informatiche (IT) e di comunicazione che possa compromettere la confidenzialità, disponibilità e integrità dei dati o dei servizi. Il danneggiamento delle tecnologie operative (OT) può portare all'interruzione dell'attività, al guasto delle infrastrutture o al danneggiamento fisico di beni e persone. Questi possono essere causati sia da disastri naturali che da azioni umane (ad esempio, *cybercrime*) e sono caratterizzati da forte interdipendenza, possibilità di eventi estremi, forte incertezza e rischio di cambiamento». Su una simile lunghezza d'onda si pone l'Institute of Risk Management (IRM) che definisce il *cyber risk* come "ogni rischio di perdita finanziaria, interruzione di servizio o danneggiamento della reputazione di un'organizzazione, derivante da un qualche tipo di guasto dei propri sistemi informatici".

Contributi a riguardo arrivano anche da altre organizzazioni internazionali riconosciute. Ad esempio, sia il World Economic Forum (2012) che il Financial Stability Board nel *Cyber Lexicon* (2018) definiscono il *cyber risk* come "la combinazione tra la probabilità di realizzazione di un *cyber incident* e il relativo impatto", ove per *cyber incident* si intende "un evento che metta in pericolo la sicurezza di un sistema informatico o delle informazioni che il sistema processa, conserva o trasmette, oppure che violi le norme e le procedure di sicurezza, sia che tale evento derivi da attività malevola o meno".

[1] Strupczewski (2021).

Com'è evidente, non vi è un consenso generale su quali siano gli elementi fondamentali per definire il rischio IT; autori diversi si focalizzano su aspetti diversi del fenomeno. Ed in effetti pare corretto affermare che *cyber risk* sia più un 'termine ombrello' sotto il quale è racchiusa una gamma di rischi diversi, risultanti da una disfunzione o da una violazione dei sistemi informatici[2]. Strupczewski (2021) propone un sistema di classificazione di queste differenti definizioni, riconoscendo 3 "elementi chiave" dei *cyber events*: le fonti, i 'bersagli' e le 'conseguenze'*; s*eguendo questa tripartizione è possibile discernere gli elementi che li caratterizzano.

Innanzitutto, con riguardo alle fonti, una prima classificazione vede distinte le **fonti interne** alle organizzazioni e le **fonti esterne** ad esse; è tuttavia possibile distinguere anche le **fonti accidentali** (e.g. danneggiamento fisico di server dovuto a un allagamento) dalle **fonti artificiali** (*man-made*), come ad esempio un attacco hacker[3].

Proseguendo con i 'bersagli' o, in senso più ampio gli 'oggetti' su cui può manifestarsi il rischio IT, secondo la Bank of International Settlements (2016) informazioni sensibili, risorse di telecomunicazione e sistemi informatici sono le categorie di asset aziendali più a rischio.

Infine, le conseguenze di un *cyber incident,* indipendentemente dalla natura dello stesso, possono essere molteplici: perdita finanziaria, interruzione temporanea dell'attività, furto o perdita di informazioni e dati sensibili. A questi effetti, che possiamo definire diretti, si associa poi una serie di effetti secondari o indiretti che oscillano dai danni reputazionali alle spese legali, oltre che ai possibili costi opportunità da danno emergente e lucro cessante[4].

1.2 Classificazione dei *cyber incident*

Diversi autori in letteratura hanno sottolineato l'importanza di introdurre un sistema di classificazione degli incidenti *cyber* adeguato e standardizzato, sia in ottica regolatoria che per supportare l'attività di gestione e mitigazione del rischio da parte delle singole organizzazioni; un sistema di questo tipo, ad esempio, faciliterebbe la costruzione di database e la condivisione di informazioni, l'identificazione di minacce emergenti e prassi comuni dei cybercriminali, oltre che la collaborazione tra autorità di regolazione[5].

Una prima categoria che è possibile individuare, come accennato in precedenza, è quella degli eventi accidentali. Questi eventi, che possono avere origine interna o esterna all'impresa, sono il risultato di azioni che non prefigurano un intento malizioso, oppure semplicemente hanno origine naturale; essi danno luogo ad un rischio puramente operativo, idiosincratico,

[2] Aldasoro et al. (2021)
[3] Eling and Schnell (2016)
[4] Curti et al. (2019)
[5] Curti et al. (2019)

connaturato all'attività stessa d'impresa, ed è dunque possibile prevenirli e gestirli adottando procedure e sistemi di controllo adeguati.

Una seconda categoria, ben più vasta, è quella degli eventi intenzionali o *man-made.* Si tratta di quegli eventi che originano dall'intento di ledere all'impresa, e anch'essi possono generarsi dall'interno (*insider threat*) o dall'esterno (*cybercrime*). I responsabili di questi attacchi (definiti *threat actors* nell'ambito della *cybersecurity*[6]), le loro motivazioni, i mezzi e le strategie impiegate possono essere molteplici, e saperli riconoscere e distinguere può essere un importante vantaggio sia in fase di prevenzione che in fase di contenimento dei danni. Un semplice esempio di sistema di classificazione dei *cyber incident* è proposto nella Tabella 1.

Ad esempio, un cybercriminale o un *insider* interessato al guadagno personale è probabile che abbia come obiettivo il furto diretto di fondi o di dati sensibili da rivendere in un secondo momento; al contrario, un gruppo di c.d. hacktivisti è spinto da motivazioni ideologiche e di giustizia sociale, ed è più incline a perseguire strategie idonee a "lanciare un messaggio" come un attacco DDoS (*Distributed Denial of Service*) che blocchi l'operatività dell'impresa target.

Threat Actor	**Motivazioni e scopi**
Cybercriminali *Insiders* "Hacktivisti" Terroristi Nazioni ostili	Profitto personale Ideologia e giustizia sociale Motivazioni geopolitiche Spionaggio industriale Incidente non intenzionale
Metodi di attacco o cause dell'incidente	**Conseguenze dell'evento**
Malware e Ransomware Attacchi DoS e DDoS *Phishing* *Skimming* Password *Cracking* Attacchi "*man-in-the-middle"* Errore Umano Avarie tecniche	Furto di fondi Frode Violazione, compromissione o furto di dati sensibili o dati industriali Costi reputazionali Interruzione temporanea dell'operatività

Tabella 1: Semplice sistema di classificazione dei cyber incident. Fonte: Aldasoro et al. (2020).

[6] Ad esempio, Center for Internet Security, "Cyber Threat Actors", https://www.cisecurity.org/spotlight/cybersecurity-spotlight-cyber-threat-actors/

1.2.1 Metodi di attacco più utilizzati

In questo paragrafo è proposta una breve spiegazione dei metodi di attacco più frequentemente utilizzati dai *threat actor*.

Malware. Abbreviazione di "Malicious Software", il termine Malware si riferisce ad una categoria di programmi ideati per danneggiare o trarre vantaggio da qualsiasi dispositivo programmabile, server o anche reti; sono usati dai cybercriminali principalmente per estorcere dati o informazioni da utilizzare come leva per conseguire un profitto, come password o dati finanziari, ma anche talvolta a fini di spionaggio. Esistono molte tipologie di malware, che utilizzano strategie di attacco diverse; tra le più redditizie e di conseguenza più diffuse vi è sicuramente il *ransomware*. Si tratta di un malware che si auto-installa su un dispositivo e ne cripta i file richiedendo poi successivamente un pagamento per ripristinare l'accesso ai dati o addirittura minacciando di pubblicare i dati personali della vittima se le richieste non sono soddisfatte. La società di sicurezza informatica Sophos ha stimato che nel 2020 il riscatto medio pagato dalle vittime di *ransomware* è stato di circa $170,404, mentre il costo complessivo medio di un attacco (comprensivo quindi di costi opportunità, costi del personale, riscatto pagato, etc.) è stato di circa $1,85 milioni[7]. Il *ransomware* è tra i metodi preferiti dai cybercriminali, soprattutto perché i pagamenti vengono richiesti attraverso metodi non tracciabili, principalmente criptovalute.

Attacchi DoS e DDoS. Un attacco *Denial-of-Service* (DoS) o *Distributed-Denial-of-Service* (DDoS) si verifica quando gli utenti legittimi non sono in grado di accedere ai sistemi informatici, ai dispositivi o alle risorse di rete a causa dell'azione di un *cyber threat actor*, il quale inonda l'*host* obiettivo con pacchetti di richieste fino al punto in cui quest'ultimo non è più in grado di gestire il traffico e si blocca; si parla poi di attacco DoS distribuito (DDoS) quando il traffico illegittimo arriva da più dispositivi che cooperano (c.d. *botnet*) per attaccare un singolo obiettivo. Questi attacchi hanno l'effetto di bloccare, per un determinato lasso temporale[8], l'accesso ai servizi dell'host, come ad esempio le e-mail, siti internet o gli account online (come avviene nel caso del settore bancario, ad esempio). Secondo un report della società F5[9], tra il 2020 e il 2021 i settori più colpiti da questo tipo di attacchi sono stati quello tecnologico, delle telecomunicazioni e il settore finanziario.

---

[7] Fonte: Sophos, "The State of Ransomware 2021", https://www.sophos.com/en-us/medialibrary/pdfs/whitepaper/sophos-state-of-ransomware-2021-wp.pdf?cmp=120469

[8] La durata media di un attacco DDoS è stimata intorno alle 3 ore, fonte: Kaspersky, https://securelist.com/ddos-attacks-in-q3-2021/104796/

[9] https://www.f5.com/labs/articles/threat-intelligence/ddos-attack-trends-for-2020

***Phishing*.** Il *phishing* è un noto metodo di truffa attraverso il quale l'autore inganna le vittime inducendole a fornire dati sensibili, come password o dati bancari. Lo strumento più utilizzato per il phishing sono le e-mail, attraverso le quali il cybercriminale si finge un'istituzione affidabile convincendo le vittime a inserire i propri dati personali in un sito web truffaldino identico a quello ufficiale, oppure alle quali è allegato un malware che carpisce le informazioni direttamente dal dispositivo. Un tipo di phishing particolarmente pericoloso per le imprese è il *Business Email Compromise* (BEC), nel quale il truffatore si appropria di una e-mail aziendale attraverso la quale si accattiva la fiducia delle proprie vittime. Secondo il report annuale 2021 di IBM questo tipo di violazione è il più costoso con un costo medio di circa $5 milioni[10].

Attacchi MITM. I "*man-in-the-middle"* sono attacchi informatici in cui un terzo si inserisce clandestinamente nella comunicazione in corso tra due parti ignare. Esistono diverse varianti di questo tipo di attacchi, che si possono verificare banalmente anche tramite connessioni Wi-Fi non protette o attraverso l'appropriazione abusiva di indirizzi e-mail; tutte hanno come risultato finale la sottrazione di dati sensibili come credenziali di accesso o numeri di carte di credito. Tipicamente, le vittime di un MITM sono utenti di applicazioni finanziarie (ad esempio, l'applicazione della propria banca) o siti di *e-commerce*.

1.3 *Cyber Risk*: un rischio operativo e sistemico

Alla luce di quanto detto, appare dunque evidente che il *cyber risk* si configuri come un (ampio) sottoinsieme della categoria dei rischi operativi sostenuti da un'impresa, e del resto questa è la tradizionale visione delle autorità di regolazione[11]. I rischi operativi sono definiti dal Comitato di Basilea come «i rischi di perdite derivanti dall'inadeguatezza o dal cattivo funzionamento di procedure, risorse umane e sistemi interni, oppure da eventi esogeni»[12]. Si tratta di una categoria di rischio che presenta una serie di peculiarità rispetto a quelle 'tradizionali', come il rischio di credito o il rischio di mercato.

Innanzitutto, il rischio operativo è un rischio puro, non speculativo, che non dà luogo a possibili guadagni, ma solo a possibili perdite e che è legato a cause accidentali non prevedibili. Ciò determina anche il fatto che per il rischio operativo non valga il principio, generalmente valido per i rischi finanziari, per il quale a un maggiore rischio è associato un maggiore guadagno atteso: non è infatti ragionevole assumere che un'impresa con un profilo di rischio operativo più alto possa attendersi profitti più alti.

---

[10] Fonte: IBM, "Cost of a Data Breach 2021", https://www.ibm.com/it-it/security/data-breach
[11] Kopp et al. (2017)
[12] È opportuno evidenziare che dalla definizione sono esplicitamente esclusi i rischi reputazionali e strategici, mentre è incluso il rischio legale.

Altra peculiarità del rischio operativo è quella di presentare una sorta di "bimodalità", ove ad incidenti molto frequenti sono associate piccole perdite, mentre ad eventi molto rari sono associate perdite di grande severità (*high frequency low impact* vs *low frequency high impact*). Inoltre, esso non è facilmente misurabile, in quanto derivante da una serie di concause molto diverse tra loro che spesso le imprese faticano a identificare e comprendere appieno, con conseguenti difficoltà anche per la raccolta dei dati necessari ad un'efficace valutazione. Infine, altro importante elemento di distinzione è la scarsità di efficaci meccanismi di mitigazione e l'assenza di un mercato secondario liquido che ne consenta il trasferimento ad una controparte come accade, ad esempio, per il rischio di mercato o per il rischio di controparte.

Al fine di riconoscere gli eventi che generano delle perdite di natura operativa, nell'ambito del *framework* di Basilea II, il Comitato ha introdotto un sistema di classificazione degli stessi caratterizzato da sette *Event Type Categories*; Curti et al. (2019), ad esempio, includono tale sistema come ulteriore elemento di categorizzazione dei *cyber incidents,* a riprova della contiguità concettuale tra rischio operativo e *cyber.*

Quanto detto sinora è ovviamente applicabile anche al sottoinsieme del *cyber risk*, che se possibile presenta ulteriori difficoltà di natura tecnica, come ad esempio nella costruzione di basi di dati complete ed esaustive in assenza di uno standard comune di rilevazione. Organizzazioni che sperimentano un *cyber incident* potrebbero essere restìe a divulgarlo per non ledere alla propria reputazione o potrebbero non avere i mezzi o le competenze tecniche necessarie a rilevarlo; per di più, anche qualora si riesca a costruire un dataset sufficientemente ampio e ben strutturato, le informazioni derivabili da esso potrebbero presto perdere di attendibilità: si tratta infatti di un ambito in rapida evoluzione a causa della incessante innovazione tecnologica.

Il *cyber risk*, tuttavia, presenta anche peculiarità difficilmente estendibili alla macrocategoria di riferimento: un attacco informatico mirato a una grande banca, ad esempio, ha potenzialmente una risonanza mediatica molto elevata, a cui si associano costi reputazionali alti, ma che sono tradizionalmente esclusi dalla definizione di rischio operativo. In questo senso è esemplificativo lo studio di Eling and Wirfs (2018) che utilizzano un database di perdite 'operative' e ne estraggono quelle riconducibili a *cyber events*, mettendole a confronto con il resto del campione: i risultati dei test statistici effettuati dagli autori mostrano che i due sotto-campioni *cyber* e *non-cyber* sembrano provenire da popolazioni differenti, e anche il Value-at-Risk[13] calcolato è più alto per le perdite *cyber-related.*

[13] Il concetto di *Value-at-Risk* (VaR) è illustrato nel Capitolo 2.

Altro punto cruciale è che il *cyber risk*, specie se inteso come 'vulnerabilità' agli attacchi esterni, ha il potenziale per trasformarsi da rischio idiosincratico in rischio sistemico, attraverso noti effetti contagio. La vulnerabilità di una singola componente è infatti sufficiente per esporre tutto il sistema: esemplificativa è la nota vicenda del *ransomware* WannaCry che nel maggio 2017 arrivò ad infettare oltre 200.000 computer in 150 paesi, bloccando l'operatività di molte istituzioni pubbliche, banche, imprese e addirittura ospedali e richiedendo il pagamento di un riscatto per liberare i dati criptati dal virus.

È dunque nell'ottica di scongiurare gravi e generalizzate conseguenze per l'economia mondiale (e non solo) che vanno letti i continui sforzi, da parte di molteplici istituzioni, di fissare degli standard comuni minimi di *cybersecurity* e *cyber-resilience*[14].

1.4 Il Cyber Risk per il sistema finanziario

Il sistema finanziario è uno dei principali canali attraverso cui il *cyber risk* può tramutarsi da rischio idiosincratico in rischio sistemico. Per la natura stessa dell'attività svolta le banche e gli intermediari finanziari in generale risultano tra le imprese più esposte al *cyber risk.* Il settore finanziario è infatti fortemente dipendente dal corretto funzionamento dei sistemi informatici e comunicativi e dei processi interni attraverso i quali si concretizzano materialmente le attività aziendali principali, come quella d'intermediazione o d'investimento; inoltre, detenendo grosse quantità di dati sensibili della propria clientela e movimentando imponenti volumi di denaro, gli istituti finanziari rappresentano un target troppo proficuo per i cybercriminali. Nish et al. (2020) fanno notare inoltre che le imprese finanziarie, più di imprese in altri settori, tendono a fare affidamento su infrastrutture e software datati, spesso 'uniti' come conseguenza di operazioni di fusione e acquisizione (*legacy infrastructure*) e che questo rappresenta una criticità del settore, che risulta più facilmente vulnerabile.

Se da un lato, comunque, è ragionevole pensare che errori umani o disfunzioni accidentali non siano significativamente più frequenti per le imprese finanziarie piuttosto che per imprese di altri settori, lo stesso non si può dire per quanto riguarda le minacce provenienti dall'esterno, in particolare sottoforma di *cybercrime* o *cyber terrorism*. La frequenza degli attacchi informatici indirizzati a istituti finanziari segue un trend crescente da anni; tra gli eventi con maggior risonanza nell'ultimo decennio ricordiamo[15]:

[14] Per un approfondimento sul *cyber risk* inteso come rischio sistemico si rimanda al report "Systemic cyber risk" del Comitato Europeo per il Rischio Sistemico (CERS).
[15] Per una *timeline* completa dei più importanti *cyber incident* dell'ultimo decennio, si rimanda a https://carnegieendowment.org/specialprojects/protectingfinancialstability/timeline

- **Attacchi DDoS a sei banche americane.** Nel settembre 2012, sei tra i più importanti istituti finanziari americani (Bank of America, JP Morgan Chase, Citigroup, U.S. Bank, Wells Fargo e PNC) furono colpite da un attacco DDoS (*Distributed Denial of Service*) che impedì alla clientela l'accesso ai servizi online per alcuni giorni[16].
- **Carbanak.** Nel 2014 fu scoperto il gruppo denominato Carbanak, composto da hacker altamente qualificati, capaci di penetrare nei network interni delle banche attraverso tecniche di *phishing* senza essere rilevati per mesi. Si stima che il gruppo sia riuscito a rubare una cifra che si aggira intorno al miliardo di dollari, attraverso l'hacking del sistema di comunicazione SWIFT o degli ATM[17].
- **JP Morgan *Data Breach.*** Sempre nel 2014, JP Morgan fu vittima di uno dei maggiori *data breach* della storia: furono infatti sottratti dati (non sensibili) di circa 83 milioni di clienti della banca[18].
- **Cyber-Rapina alla Banca Centrale del Bangladesh.** All'inizio del 2016 fu reso noto che un gruppo di hacker aveva cercato di rubare circa un miliardo di dollari dalla Banca Centrale del Bangladesh (BCB), inviando 35 ordini fraudolenti di trasferimento fondi dal conto che la BCB deteneva presso la Federal Reserve di New York verso conti sparsi in tutto il mondo attraverso il sistema SWIFT. Dei 35 ordini solo 5 ebbero successo, per un ammontare di circa 81 milioni di dollari complessivi sottratti in maniera illegale. L'evento ebbe una forte risonanza mediatica, fungendo da campanello d'allarme per l'intero settore bancario e anche risvolti di natura politica, in quanto la Corea del Nord fu sospettata di essere il mandante dietro agli attacchi[19].
- **Binance Ransomware.** Nell'Agosto del 2019, l'*exchange* di criptovalute Binance fu vittima di un *ransomware*. Gli hacker domandarono un riscatto di circa 3,5 milioni di dollari in bitcoin in cambio di un database sottratto alla società contenente le informazioni personali di circa diecimila clienti.
- **Glitch di Target2.** Nell'ottobre 2020, a causa di un *glitch*, fu interrotto per circa 11 ore il servizio del principale sistema dei pagamenti attivo nell'Eurozona, Target2, gestito dalla BCE[20].

[16] Fonte: New York Times, https://www.nytimes.com/2012/10/01/business/cyberattacks-on-6-american-banks-frustrate-customers.html
[17] Fonte: Kaspersky, https://www.kaspersky.com/blog/billion-dollar-apt-carbanak/7519/
[18] Fonte: Reuters, https://www.reuters.com/article/us-jpmorgan-cybersecurity-idUSKCN0HR23T20141003
[19] Fonte: BBC News, https://www.bbc.com/news/stories-57520169
[20] Fonte: The Wall Street Journal, https://www.wsj.com/articles/europes-core-payments-network-disrupted-by-technical-malfunction-11603899692

Come mostrato in Aldasoro et al. (2020b), la frequenza degli attacchi diretti al settore finanziario è cresciuta fortemente fino al 2016 ed è successivamente diminuita leggermente; un drammatico rialzo si è avuto però durante la pandemia da Covid-19, a causa principalmente della necessità del lavoro in remoto che ha aumentato le vulnerabilità dei sistemi aziendali. Un aumento di circa il 238% è stato registrato tra i mesi di Febbraio e Aprile 2020, con l'80% di istituti finanziari che hanno riportato un aumento degli attacchi[21]; proprio nell'aprile 2020 il Financial Stability Board avvertiva in un comunicato stampa che «un *cyber incident* di entità rilevante, se non propriamente controllato, potrebbe seriamente intaccare i sistemi finanziari, comprese infrastrutture finanziarie fondamentali, e portare a più ampie conseguenze dal punto di vista della stabilità finanziaria»[22].

Sebbene sia tra i più colpiti, comunque, il settore finanziario sembra sperimentare perdite relativamente più basse: circa $1.7 milioni per evento a fronte di una media di $2.6 milioni per tutti i settori[23]. Un risultato simile è illustrato anche in Eling and Wirfs (2019), nel cui studio è mostrato che circa il 76% dei *cyber incidents* registrati riguarda l'industria finanziaria, ma con una perdita media più bassa: $30,57 milioni a fronte di una media di $84,11 milioni per gli altri settori[24].

Come si spiega ciò? McKinsey (2021) nel resoconto di un sondaggio condotto presso più di 100 imprese e organizzazioni ha riscontrato che il settore finanziario è tra i primi 3 per livello medio di sviluppo della *cybersecurity*[25]; questo è un risultato che in effetti non sorprende più di tanto, in quanto si tratta di un settore che da un lato dipende fortemente dalla fiducia dei consumatori dovendone trattare e conservare i dati sensibili oltre che i risparmi, e che dall'altro si trova sotto la lente d'ingrandimento delle autorità di regolazione e di controllo.

1.5 Il panorama legislativo europeo: leggi, direttive e regolamenti per il settore finanziario europeo in materia di *cybersecurity* e *cyber risk management*

La legislazione europea in materia di cybersecurity per il settore finanziario è abbastanza complessa e strutturata su più livelli; non esiste una legislazione unica, quanto più una moltitudine di atti (leggi, regolamenti, direttive, linee guida) che non in tutti i casi hanno valore legalmente vincolante. A questo si aggiunge il fatto che gli standard in materia non sono

[21] Fonte: Allianz, "Financial Services Risk Trends" (2021)
[22] FSB, https://www.fsb.org/2020/10/fsb-encourages-use-of-cyber-incident-response-and-recovery-toolkit/
[23] Aldasoro et al. (2020b)
[24] Gli stessi autori sottolineano come i risultati da loro ottenuti sono più alti rispetto a quelli di altri studi; una parziale spiegazione è attribuibile alla differenza di database, ove quello da loro utilizzato non include perdite inferiori ai $100,000 oltre ad includere anche *incidents* registrati al di fuori degli USA.
[25] McKinsey, https://www.mckinsey.com/business-functions/risk-and-resilience/our-insights/organizational-cyber-maturity-a-survey-of-industries

uniformi, ma differenziati tra i vari comparti che compongono il settore finanziario (comparto bancario, comparto assicurativo, mercati finanziari, settore dei pagamenti etc.) e spesso tendono anche a sovrapporsi, com'è il caso per gli istituti di credito che, operando in diversi ambiti, devono tener conto di regolamentazioni diverse.

Ad un primo livello generale si pongono due direttive di ampio respiro, non specifiche di settore, emanate nel 2016: la *Directive on Security of Network and Information Systems* (Direttiva NIS) e il *General Data Protection Regulation* (GDPR). La direttiva NIS, recepita in Italia con il d.lgs. 18 maggio 2018, n.65 e attualmente in fase di revisione da parte dell'Unione, mira a definire le misure necessarie per ottenere un adeguato livello di sicurezza delle reti e dei sistemi informatici; essa definisce due categorie di enti, gli **operatori di servizi essenziali (OSE)** e i **fornitori di servizi digitali (FSD)**, a cui si applicano le misure previste. In particolare, il gruppo degli OSE è composto da soggetti, pubblici o privati, operanti in sette settori definiti essenziali per il funzionamento della società e dell'economia; due di questi settori sono, appunto, il comparto bancario e le infrastrutture per i mercati finanziari. Tra le misure più rilevanti contenute nella direttiva vi sono[26]: a) obbligo di adozione di misure tecniche ed organizzative adeguate e proporzionate alla gestione dei rischi e a prevenire e minimizzare l'impatto degli incidenti; b) obbligo di notificare, senza ritardo, gli incidenti con impatto rilevante sulla continuità e sulla fornitura dei servizi. Il GDPR, che a differenza della NIS è operante a livello comunitario senza bisogno di ricezione nei singoli ordinamenti nazionali, mira a fornire una legislazione standard unica in materia di protezione dei dati e a dare un maggiore controllo ai cittadini su come i propri dati vengono utilizzati; in questo senso, si applica a tutti quegli enti che processano o controllano dati personali e che operano nel territorio dell'Unione. Esso prevede, tra le altre cose, l'obbligo di riferire eventuali violazioni dei dati personali e multe fino a €20 milioni per la non osservanza delle disposizioni. L'impatto di questo regolamento sul settore finanziario è stato alto, contribuendo ad alzarne il livello di *cybersecurity* e fornendo una base istituzionale per la raccolta di dati sui *data breach.*

A queste fonti normative di ampio respiro si affianca poi un *corpus* di leggi, direttive e regolamenti specifici per ogni comparto del settore finanziario, di cui un'ampia trattazione è offerta in Krüger and Brauchle (2021), mentre in questa sede ci concentreremo sugli istituti di credito. Per questi enti, infatti, l'insieme di norme esistente è particolarmente complesso e stratificato in quanto trattasi di organizzazioni dall'elevato grado di digitalizzazione, di rilevante importanza sistemica e che offrono diversi servizi oltre a quelli legati alla tradizionale attività bancaria, in particolare servizi di pagamento e d'investimento. La frammentarietà di

[26] Per maggiori informazioni si rimanda a https://www.sicurezzanazionale.gov.it/sisr.nsf/archivio-notizie/cyber-la-nis-entra-in-vigore-litalia-si-rafforza-e-fa-rete-con-lue.html

questo quadro legislativo ha portato più autori a caldeggiare l'introduzione di un sistema più unitario, che riduca le sovrapposizioni e crei meno incertezza per le parti in causa[27].

Un primo insieme di disposizioni in materia proviene dal pacchetto CRR/CRD IV[28], provvedimenti attuativi del Terzo Accordo di Basilea (Basilea III) nell'ambito dell'Unione in cui, riflettendo la visione del Comitato, la regolamentazione del *cyber risk* è trattata in maniera implicita in quanto considerato sottoinsieme del rischio operativo. Si tratta di disposizioni di natura prudenziale il cui primo obiettivo è quello di indicare dei requisiti di capitale minimi che gli intermediari devono mantenere per far fronte ai vari rischi insiti nella propria attività (tra cui, appunto, il rischio operativo); in secondo luogo, il pacchetto normativo prevede anche l'obbligo di istituire e implementare processi di valutazione dell'esposizione al rischio operativo e a preparare piani d'azioni per garantire la continuità aziendale anche in situazioni d'emergenza[29]. A corredo di questo primo *corpus* normativo, la EBA ha emanato anche un set di linee guida specifiche per la gestione del *cyber risk*[30].

Tra gli altri provvedimenti rilevanti per gli istituti di credito si rilevano la direttiva PSD2, rivolta ai fornitori di servizi di pagamento, e la MIFID II, direttiva rivolta ai fornitori di servizi d'investimento. In particolare,

- la *Revised Payment Services Directive* (PSD2) prevede, come condizione per il rilascio dell'autorizzazione a operare nel settore dei pagamenti, un alto livello di sicurezza informatica e di protezione dei dati, oltre a specifiche previsioni per quanto riguarda l'*outsourcing* di funzioni operative, tra cui anche la gestione dei sistemi informatici;
- la *Markets in Financial Instruments Directive 2* (MIFID II), similmente, prevede l'obbligo a carico dei soggetti interessati di "adottare misure, risorse, procedure e sistemi adeguati e proporzionati", di "stabilire meccanismi di controllo interno, procedure di valutazione del rischio oltre che misure di salvaguarda dei sistemi informatici" e di "utilizzare meccanismi di sicurezza tali da minimizzare il rischio di violazione dei dati.

Infine, a chiusura del sistema legislativo rilevante è importante citare anche il *Cyber Incident Reporting Framework* che si applica a tutte le banche di rilevanza sistemica, poste direttamente sotto l'egida della BCE; le disposizioni contenute nel *framework* pongono

[27] Callies and Baumgarten (2020)
[28] Capital Requirements Regulation (CRR, Direttiva 2013/36/UE) e Capital Requirements Directive IV (CRD IV, Regolamento n.575/2013 UE)
[29] Krüger and Brauchle (2021)
[30] EBA, Guidelines on ICT and security risk management, https://www.eba.europa.eu/regulation-and-policy/internal-governance/guidelines-on-ict-and-security-risk-management

l’obbligo, in capo a queste banche, di riportare direttamente alla BCE i *cyber incident* di maggior rilevanza.

Capitolo 2

Metodologia

2.1 Introduzione al capitolo

In questo capitolo sarà illustrata la metodologia scelta per l'analisi con le relative fondamenta teoriche e saranno introdotti alcuni concetti fondamentali relativi alle misure di rischio. Come anticipato nel primo capitolo, si è scelto di approcciare al problema da una prospettiva di puro *risk management*, tralasciando gli aspetti strettamente tecnici in materia di *cybersecurity* e gli aspetti organizzativi relativi, ad esempio, ai processi decisionali e alle attribuzioni di responsabilità aziendale. Lo scopo dell'analisi sarà infatti quello di fornire una singola misura, quella del *Value-at-Risk* relativo al rischio informatico, e di discuterne i potenziali usi.

A questo fine si è scelto di utilizzare gli strumenti adottati comunemente in materia di quantificazione del rischio operativo, e in particolare si farà riferimento al metodo denominato *Loss Distribution Approach* (LDA), tipico delle scienze attuariali e già utilizzato da diversi autori per scopi simili[31]. Questo approccio è infatti tra i più utilizzati dagli istituti finanziari che hanno scelto di ricorrere all'*Advanced Measurement Approach*, introdotto dal Nuovo accordo sul Capitale delle Banche (c.d. Basilea II), per la misurazione del rischio operativo[32]. In estrema sintesi, il LDA consiste nel trovare e combinare la distribuzione di frequenza degli eventi che causano perdite con la distribuzione di severità di queste ultime; a partire da questa combinazione, attraverso tecniche di simulazione come il metodo Montecarlo, viene generato un numero predeterminato di scenari sulla base dei quali si otterrà la distribuzione totale delle perdite in un dato orizzonte temporale, dalla quale a sua volta sarà possibile ricavare gli indicatori di rischio cercati. Com'è evidente si tratta di un metodo particolarmente oneroso sia in termini di *input*, in quanto richiede dati sia sulla frequenza che sulla severità delle perdite, che di calcolo, in quanto richiede di generare un elevato numero di scenari; sfortunatamente, nell'ambito del *cyber risk*, non sono disponibili dataset pubblici affidabili e completi e ciò rappresenta un grosso ostacolo per la ricerca in materia.

[31] Ad esempio, Bouveret (2018) o Eling and Wirfs (2015).

[32] Bisogna tuttavia sottolineare che il BCBS, nei nuovi accordi che vanno sotto il nome di Basilea III e che entreranno in vigore definitivamente nel 2023, ha scelto di semplificare il metodo di misurazione del requisito di capitale relativo al rischio operativo, introducendo un approccio unico comune a tutte le banche in luogo dei tre finora consentiti. Ciò nonostante, l'applicazione del LDA nell'ambito della gestione dei rischi interna resta comunque valida.

Ciò che resta da chiarire è il motivo per cui un istituto finanziario dovrebbe realizzare dei sistemi specifici per quantificare la propria esposizione al *cyber risk.* Allo stato attuale, infatti, l'attitudine generale nei confronti di questo problema è fortemente improntata verso la mera soddisfazione degli standard legislativi (*compliance*) piuttosto che verso la effettiva costruzione di un sistema sicuro ed affidabile; ed in effetti, quest'atteggiamento è parzialmente giustificato dal fatto che si tratta di una problematica nuova e non pienamente compresa dalla maggior parte delle aziende nei diversi settori. In virtù del riconosciuto potenziale trainante ed innovante, è naturale dunque pensare a quello finanziario come uno dei settori che possono portare ad un cambio di paradigma.

Gli incentivi, dunque, possono essere diversi. Innanzitutto, un indicatore puramente finanziario di questo tipo può aiutare il personale non-tecnico a comprendere l'importanza di porre in essere comportamenti e processi idonei a preservare la sicurezza dell'organizzazione[33]; può essere utile ad orientare le scelte d'investimento in sicurezza informatica da parte del management; e in ultima analisi garantire che adeguate riserve di capitale siano poste a salvaguardia dell'azienda per proteggerla da una minaccia non ancora pienamente compresa, in costante evoluzione e crescita e che rischia di rappresentare una zavorra per lo sviluppo futuro dell'intero settore.

2.2 *Risk Measures*

Come accennato, lo scopo dell'analisi sarà quello di fornire una misura concreta, sebbene indicativa, del *cyber risk*; in virtù di ciò è innanzitutto necessario introdurre questa misura, il Value-at-Risk (VaR). Sviluppato a partire dalla prima metà degli anni Ottanta per la misurazione del rischio di mercato, il concetto di VaR[34] ha da allora progressivamente acquisito importanza fino a divenire il modello prevalente nel campo della misurazione e della gestione dei rischi in generale, fino ad esser preso come riferimento anche dalle autorità di regolazione nel 1992 col primo accordo di Basilea. Gran parte di questa importanza è dovuta soprattutto alla semplicità concettuale del VaR, il quale essenzialmente punta a fornire la risposta alla domanda

> *«Qual è la perdita massima a cui si può andare incontro in un determinato periodo temporale, tale che vi sia una probabilità molto bassa che la perdita effettiva risulti superiore al valore stimato?»*

[33] È lo stesso comitato di Basilea a sottolineare l'importanza che il personale, a qualsiasi livello, abbia un'adeguata consapevolezza in termini di *risk culture* come prerequisito essenziale per la resilienza informatica degli istituti. Fonte: "Cyber-resilience: Range of Practices" BCBS (2018) https://www.bis.org/bcbs/publ/d454.pdf

[34] Il termine Value-at-Risk è utilizzato per indicare, alternativamente, sia la misura stessa che il modello concettuale che vi è alla base.

Da questo quesito è possibile desumere i due elementi principali comuni a tutti i modelli appartenenti alla famiglia VaR, e cioè

- La presenza di un orizzonte temporale fissato, che può essere più o meno breve in relazione alla tipologia di rischio che si tratta; ad esempio, la valutazione del rischio di mercato richiederà sicuramente un orizzonte temporale molto breve, tipicamente giornaliero o settimanale, mentre il rischio operativo può essere valutato su un periodo più lungo, tipicamente annuale;
- Un livello di confidenza predeterminato, cioè la probabilità con la quale siamo sicuri che la stima VaR non sarà superata dalla perdita effettiva. Naturalmente, scegliere un livello di confidenza più alto porterà anche ad una stima di valore a rischio più alto; tipicamente, il VaR viene stimato al livello di confidenza del 95, 97.5 o 99 percento. Stabilire un livello di confidenza del 99% significa che la stima effettuata sarà superata con probabilità pari solo all'1%.

Da un punto di vista formale, quanto detto è sintetizzabile nel modo che segue.

Definizione 1. Sia X la variabile casuale rappresentativa delle perdite sopportate durante un determinato periodo di tempo t (ad esempio un anno); sia poi $\alpha \in (0,1)$ il livello di confidenza prescelto. Allora, il Valore-a-Rischio sarà il quantile α della distribuzione tale che

$$VaR_\alpha(X) = \inf\{x \in \mathbb{R} : P(X > x) \leq (1-\alpha)\} = \inf\{x \in \mathbb{R} : F_X(x) \geq \alpha\} \qquad (2.1)$$

o, in altre parole, la perdita massima che sarà superata con probabilità pari solo a 1-α. Il VaR fornisce dunque il notevole vantaggio di riassumere in un singolo numero e quindi rendere facilmente comprensibili le informazioni relative ai rischi assunti da un istituto finanziario, anche per il personale non tecnico. La figura 2 illustra il concetto di VaR: da un punto di vista matematico, l'area evidenziata è pari alla probabilità che la perdita effettiva sia superiore al VaR; come si può notare si tratta di un'area molto piccola, pari a 1-α.

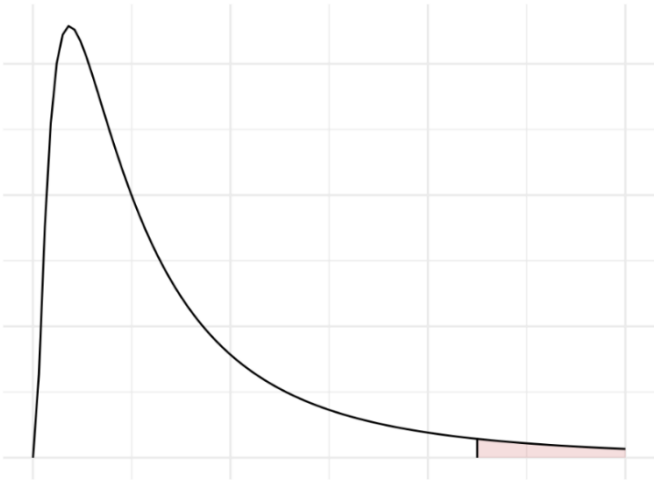

Figura 2: Illustrazione grafica del concetto di VaR. La curva in figura corrisponde alla densità di una distribuzione log-normale. Fonte: elaborazione dell'autore.

I modelli VaR costituiscono una famiglia di tecniche diverse, le quali però sono accomunate dall'impianto concettuale di base appena descritto. Per quanto riguarda la materia in esame, e cioè la quantificazione del cyber risk, i due approcci principali sono quello basato sulla distribuzione parametrica e quello basato sulla distribuzione empirica delle perdite. In breve, nel primo approccio si cerca di modellare i dati a disposizione sulla base di una distribuzione conosciuta (ad esempio la distribuzione esponenziale o la Weibull) stimandone i parametri attraverso metodi di stima come il metodo della massima verosimiglianza; nel secondo approccio (detto anche della simulazione storica) invece non si fa alcuna assunzione riguardo la forma funzionale della distribuzione, e si procede direttamente a calcolare il VaR sulla distribuzione empirica delle perdite ricavata dai dati a disposizione. Il Loss Distribution Approach seguito nell'analisi presente in questo elaborato rientra nel primo caso.

Per quanto attrattivo nella sua semplicità, il VaR presenta tuttavia anche alcune limitazioni, la più importante delle quali è costituita dal fatto che esso non è una misura di rischio sub-additiva[35]; da un punto di vista economico, ciò significa che il VaR non riesce a catturare l'effetto diversificazione che, com'è noto, fa abbassare il rischio complessivo di un portafoglio finanziario, potendosi dunque verificare situazioni in cui, indicando con A e B due portafogli distinti,

$$VaR(A+B) \geq VaR(A) + VaR(B) \qquad (2.2)$$

Questo limite si fa particolarmente gravoso considerando il fatto che il VaR viene attivamente utilizzato per stabilire i requisiti patrimoniali che le istituzioni finanziarie devono detenere in osservanza delle disposizioni prudenziali; aggregando i vari portafogli che una banca detiene si può generare una situazione in cui il requisito patrimoniale aumenti invece di

[35] La condizione di sub-additività è una delle quattro condizioni definite da Artzner et al. (1999) che una misura di rischio deve rispettare per poter essere definita "coerente" e quindi confrontabile con altre misure di rischio.

diminuire, con conseguenze negative soprattutto per la competitività della banca. Un altro limite del Value-at-Risk è che, pur indicando la massima perdita potenziale nel α % dei casi, non fornisce alcuna indicazione su ciò che avviene nel restante (1- α) %, o in altre parole non fornisce alcuna informazione sull'effettiva entità delle perdite qualora il caso peggiore dovesse effettivamente verificarsi. Una misura di rischio che supera queste limitazioni è il *Conditional Value-at-Risk* (CVaR), detto anche *Expected Shortfall* (ES); esso è strettamente correlato al VaR, e come quest'ultimo è funzione dell'orizzonte temporale e del livello di confidenza prescelti, ma rispettando la condizione di sub-additività e configurandosi dunque come una misura di rischio coerente. L'ES fornisce una risposta alla domanda

«Se la perdita reale superasse il valore stimato dal VaR, a quanto ammonterebbe la perdita attesa?»

In termini più formali:

$$ES_\alpha = E[L|L > VaR_\alpha] \qquad (2.3)$$

ossia "il valore atteso di tutte le perdite superiori al VaR"[36]. Sebbene si tratti di una misura di rischio potenzialmente migliore, l'utilizzo dell'ES può essere limitato in alcuni contesti, come ad esempio nell'ambito del rischio operativo; in linea generale, infatti, le distribuzioni di probabilità dei rischi operativi presentano code molto spesse al punto che la *expected loss* potrebbe anche non esistere (*infinite mean distributions*)[37]. Inoltre, a differenza del VaR, l'*Expected Shortfall* non si presta al *backtesting.*

2.2.1 Il *Cyber VaR*

La prima proposta organica per l'applicazione di una misura di rischio tradizionalmente finanziaria come il VaR nell'ambito della sicurezza informatica fu formulata nel 2015 nell'ambito dell'iniziativa per la *Cyber-Resilience* patrocinata dal World Economic Forum, che diede vita al *framework* concettuale denominato *Cyber Value-at-Risk* (Cy-VaR). Sebbene non fornisca un vero e proprio approccio operativo standardizzato per la quantificazione del *cyber risk*[38], la proposta del WEF contiene tuttavia delle indicazioni preziose ai fini della costruzione di un metodo completo, standardizzato e trasversale ai diversi settori dell'economia, incoraggiando poi le singole imprese a costruire i propri modelli di quantificazione interni; ogni settore e ogni azienda ha le proprie specificità, ma la comparazione tra i vari modelli è resa

[36] Resti e Sironi (2008)
[37] Cruz, Peters and Shevchenko (2015)
[38] Tra le varie proposte, la metodologia denominata FAIR™ (*Factor analysis of information risk*) è emersa come lo standard internazionale per il calcolo del Cyber Value-at-Risk. Per maggiori informazioni, https://www.fairinstitute.org/about

possibile proprio dal riferimento ad un *framework* concettuale di base condiviso. L'obiettivo ultimo della *initiative* è infatti quello di fornire un approccio unificato al problema delle *cyber threats* che, creando forte incertezza in virtù della propria natura in continua evoluzione, rappresentano uno dei più grossi ostacoli allo sviluppo economico divenuto ormai strettamente connesso a quello tecnologico. Il Cy-VaR si presta dunque ad essere più un indicatore *proxy* per l'esposizione al *cyber risk* che una stima puntuale di quest'ultimo; nondimeno, la standardizzazione attraverso i vari comparti dell'economia di un simile indicatore ne incrementerebbe sicuramente l'utilità.

Il modello concettuale del Cy-VaR richiede alle imprese, innanzitutto, di comprendere sia i fattori chiave necessari alla costruzione di un modello per il *cyber risk* che le dipendenze e le interazioni esistenti fra di essi; le componenti fondamentali sono essenzialmente tre: la **vulnerabilità** dei sistemi aziendali, gli **asset** che potenzialmente possono essere obiettivo di attacchi informatici e il **profilo del *threat actor,*** di cui un'analisi è fornita nel primo capitolo. A titolo esemplificativo, il successo di un attacco informatico è determinato dall'interazione tra la prima e la terza componente; oppure, il livello di attrattività per un *cyber* criminale di un'azienda è influenzato dal valore dei propri asset (seconda componente) e dalle tendenze più recenti riguardanti gli attacchi informatici (terza componente). Analizzare e quantificare queste tre componenti, studiandone le interazioni e le dipendenze, consentirebbe dunque di pervenire a un modello stocastico capace di quantificare il livello di *cyber risk* cui ogni singola impresa è esposta e di riassumerlo in una misura che permetta di affermare con un elevato grado di confidenza che un *cyber event* non causerà perdite superiori al valore stimato, similmente all'informazione fornita dal VaR di stampo finanziario classico.

Il Cy-VaR è dunque una misura di tipo economico che si presta ad essere utilizzata come complemento alle valutazioni strettamente tecniche sul livello di sicurezza informatica di un'impresa, e che quantificando l'impatto economico degli attacchi *cyber* aiuta anche ad orientare le scelte dei manager riguardanti gli investimenti in materia di *cybersecurity* e mitigazione del rischio (ad esempio, l'acquisto di una polizza assicurativa) o per quantificare la riduzione dell'esposizione derivante da queste scelte. Un esempio pratico di questo tipo di applicazione è fornito in Orlando (2021), in cui è proposto un indicatore di rendimento degli investimenti in sicurezza informatica aggiustato per il rischio, denominato RaROSI[39] (*risk-adjusted return on security investments*) e costruito in questo modo:

[39] Il RaROSI è un indicatore di rendimento degli investimenti basato sul ROSI, sviluppato dalla European Network and Information Security Agency (ENISA 2012) e a sua volta basato sul ben noto indicatore ROI (*return on investment*) utilizzato per giudicare la profittabilità di un investimento. Un altro indicatore basato sul ROI e proposto sempre in Orlando (2021) è il Cyber-RAROC.

$$RaROSI_{\alpha} = \frac{\Delta U[L] - I_0}{I_0} \quad (2.4)$$

dove

$$\Delta U[L] = E[L] - mCyVaR(\alpha)$$

- E[L] è la perdita attesa mentre mCyVaR è il Cyber Value-at-Risk, rappresentativo della perdita peggiore possibile, mitigato dall'investimento in sicurezza informatica;
- ΔU[L] rappresenta dunque la riduzione della perdita attesa dovuta all'investimento in sicurezza;
- I_0 rappresenta invece il costo dell'investimento in sicurezza

Gli obiettivi che il Cy-VaR si propone di raggiungere possono essere quindi riassunti in due punti principali: il primo è quello di fornire ai professionisti di sicurezza informatica un modo per esprimere il rischio informatico in un linguaggio di più ampia comprensione, quello economico; il secondo obiettivo è quello di aiutare le imprese a compiere scelte più razionali e redditizie nella protezione dei propri asset. Il raggiungimento effettivo di questi obiettivi è però ostacolato da alcune limitazioni, prima fra tutte la già evidenziata scarsità di dati reali sulla frequenza dei *cyber event* e la severità delle perdite ad essi associate. Un altro grosso limite è rappresentato dall'assenza di un *framework* standard per la valutazione della maturità dei sistemi di sicurezza informatici: come evidenziato da Buith e Spataru (2015) il numero di incidenti che un ente può patire dipende in parte anche dal livello di maturità della *cybersecurity*, e l'assenza di una misura standardizzata per quest'ultimo introduce un elemento di soggettività che limita l'applicabilità dei modelli di tipo Cy-VaR.

2.3 L'approccio di Basilea

Gli strumenti utilizzati per l'analisi presente in questa ricerca, come già anticipato, sono gli stessi comunemente utilizzati per trattare il rischio operativo, di cui il *cyber risk* è un sottoinsieme pur con le proprie peculiarità fin qui analizzate. Il rischio operativo è una tipologia di rischio a sua volta molto diversa rispetto ai rischi più "tradizionali" dell'attività bancaria, e appare dunque opportuno allargare leggermente il focus per analizzare il modo in cui esso viene regolamentato e gestito.

Il concetto di rischio operativo non è recente; ciò che è cambiato è in realtà il modo in cui esso viene percepito e trattato. Infatti, se fino agli ultimi anni del secolo scorso le perdite di natura operativa erano rare e perfettamente sostenibili, specialmente se comparate a quelle relative al rischio di credito o al rischio di mercato, uno degli effetti collaterali della deregolamentazione e della (conseguente) globalizzazione che ha interessato il settore bancario

negli ultimi vent'anni è stato proprio quello di una maggiore esposizione al rischio operativo[40]. Come risultato di questo processo, il Comitato di Basilea inserì una specifica regolamentazione del rischio operativo nell'ambito della riforma che va sotto il nome di Basilea II, iniziata nel 1998 e finalizzata nel 2006. Nell'ambito del *framework* stabilito da questi nuovi accordi fu prevista infatti una specifica riserva di capitale (riserva prudenziale) atta a proteggere le banche dalle perdite di natura operativa, similmente a quanto avveniva già per il rischio di mercato e per quello di credito; contestualmente, il Comitato suggerì anche tre diversi approcci perseguibili dai singoli istituti per il calcolo di questa riserva obbligatoria, caratterizzati da un crescente grado di complessità. Questi sono:

- Basic Indicator Approach (BIA)
- Traditional Standardised Approach (TSA)
- Advanced Measurement Approaches (AMA)

I primi due sono approcci di tipo *top-down*, ossia determinano un costo totale per il rischio operativo partendo da dati passati e/o settoriali, senza fare tuttavia distinzione tra le varie tipologie di *risk event* o tra i rami d'azienda che li hanno generati; sono approcci tipicamente semplici da implementare e che richiedono pochi input, ma che oltre ad essere poco sofisticati (generando dunque il rischio di sovrastimare o sottostimare la riserva obbligatoria) non forniscono informazioni aggiuntive, ad esempio, sui punti deboli della banca. Questi approcci prevedono il calcolo della riserva come proporzione di una certa variabile come, ad esempio, i ricavi o il reddito.

Gli AMAs invece costituiscono una famiglia di approcci di tipo *bottom-up*, che consistono nell'analisi del rischio operativo partendo "dal basso", ossia partendo dai dati raccolti all'interno della banca stessa, mappando ogni *risk event* in categorie specifiche e solo successivamente aggregando i dati per costruire, ad esempio, analisi di scenario o stress test. Si tratta dunque di approcci che incorporano il non banale vantaggio di spiegare i meccanismi che portano una banca ad avere una determinata esposizione di rischio, e che sono per questo motivo anche più complessi da implementare.

2.3.1 *Basic Indicator Approach*

Il primo degli approcci proposti è il BIA, sicuramente il più semplice e immediato; sotto questo approccio, il requisito patrimoniale obbligatorio è calcolato come una percentuale fissa del margine d'intermediazione lordo.

$$RP_{BIA} = \alpha \times \frac{\sum_{i=1}^{n} MI_i}{n} \qquad (2.5)$$

[40] Chernobai et al. (2007)

dove
α = coefficiente fissato; MI = margine di intermediazione lordo; n = numero di anni in cui MI è stato positivo (max fino a tre anni precedenti).

Come è possibile notare dalla formula, il calcolo è effettuato sul margine d'intermediazione lordo medio degli ultimi 3 anni, prendendo in considerazione solo gli anni in cui tale variabile è stata positiva[41]; il coefficiente α è invece fissato arbitrariamente dal Comitato al 15%. Si tratta di un metodo, come detto, molto semplice e particolarmente adatto a banche di piccole e medie dimensioni, che fu tipicamente utilizzato nelle prime fasi di implementazione di Basilea II. In quanto poco raffinato, è un approccio ovviamente inadeguato per banche di grosse dimensioni e/o di importanza sistemica a cui è appunto precluso. Lo svantaggio principale del BIA è che tende a sovrastimare il requisito di capitale necessario, oltre ovviamente a non fornire alcun tipo di utilità all'istituto al di là del mero soddisfacimento delle norme prudenziali.

2.3.2 *Traditional Standardised Approach*[42]

Il TSA è il secondo degli approcci proposti all'interno del *framework* di Basilea II, caratterizzato da un grado di sofisticazione maggiore rispetto al BIA. Esso prevede la suddivisione delle attività della banca in otto diverse linee di business e ad ognuna di esse viene attribuita una quota del margine d'intermediazione totale della banca; tale quota serve come indicatore *proxy* della dimensione relativa di ogni specifica linea di business. Ogni quota viene poi moltiplicata per un coefficiente, denominato beta, che varia a seconda della linea cui si riferisce. In formule:

$$RP_{TSA} = \frac{\sum_{j=1}^{3} \max\left\{\sum_{k=1}^{8}(MI_{jk} \times \beta_k)\,, 0\right\}}{3} \qquad (2.6)$$

Anche in questo caso, la variabile rilevante è dunque la media degli ultimi 3 anni del margine lordo d'intermediazione calcolato per ogni singola linea di business; qualora, per un dato anno, il margine lordo totale sia negativo allora in quel caso l'input al numeratore della formula sarà 0. In Tabella 2 sono indicate le 8 linee di business rilevanti con i relativi coefficienti beta.

[41] BCBS, "Calculation of RWA for operational risk: Basic Indicator Approach", Bank for International Settlements (2019)
[42] BCBS, "Calculation of RWA for operational risk: Standardised Approach", Bank for International Settlements (2019)

Business Line	Coefficiente Beta
Corporate Finance	18%
Trading and Sales – Intermediazione e Vendite	18%
Retail Banking – Servizi bancari al dettaglio	12%
Commercial Banking	15%
Payment and Settlement – Pagamenti e Liquidazioni	18%
Agency Services – Servizi di agenzia	15%
Asset Management – Gestione patrimoniale	12%
Retail Brokerage – Intermediazione mobiliare al dettaglio	12%

Tabella 2: Linee di business e relativi coefficienti TSA.
Fonte: BIS (2019), "Calculation of RWA for Operational Risk: Standardised Approach"

Sebbene si tratti di un metodo più avanzato rispetto al BIA, i vantaggi del TSA sono essenzialmente gli stessi: semplicità di calcolo e di implementazione, nessuna necessità di raccogliere dati granulari ma con il vantaggio di offrire una stima più precisa del requisito obbligatorio (abbassando, ma non eliminando, il rischio di sovrastima). Anche questo approccio è adatto principalmente per banche di piccole e medie dimensioni.

Il principale limite del TSA è che i coefficienti beta, essendo fissati e comuni a tutti gli istituti, potrebbero non riflettere le peculiarità di ogni banca e l'importanza che le diverse *business line* rivestono al loro interno: ad esempio, la banca Alfa potrebbe operare molto più pesantemente nel settore retail rispetto alla banca Omega il cui business principale è invece rappresentato dall'intermediazione. I coefficienti resterebbero gli stessi per entrambe, pur non riflettendone gli effettivi profili di rischio operativo.

Il *framework* di Basilea II prevede, oltre al TSA fin qui descritto, una sua versione alternativa denominata *Alternative Standardised Approach* (ASA), a cui le banche potevano accedere su richiesta e a discrezione delle relative autorità nazionali di supervisione. In questo approccio alternativo, il margine lordo d'intermediazione è rimpiazzato dalla voce di bilancio "prestiti e anticipazioni" (moltiplicata per un coefficiente fisso *m*) come indicatore di esposizione per due particolari linee di business: *retail* e *commercial banking.* Per queste due linee di business, il requisite di capitale è calcolato mediante la seguente formula

$$RP = \beta \times m \times L\&A$$

ove i relativi fattori β sono gli stessi previsti dalla tabella 2. Per quanto riguarda il comparto del *retail banking*, il valore L&A associato è il totale dei prestiti verso clienti retail e piccole e

medie imprese trattate come clienti retail; nel comparto *commercial banking*, tale valore è uguale al totale dei seguenti portafogli crediti: clienti *corporate*, prestiti sovrani, banche, piccole e medie imprese trattate come clienti *corporate*.

2.3.3 *Advanced Measurement Approaches*

Come anticipato, quella degli AMAs è più una famiglia di approcci diversi che un singolo metodo come il BIA o il TSA; la normativa si limita infatti solo a stabilire dei requisiti minimi di ammissibilità, lasciando alle singole banche la facoltà di utilizzare il proprio modello interno utilizzato per la misurazione del rischio operativo previa verifica della sussistenza di alcune condizioni[43] da parte delle autorità di vigilanza nazionali. Tra queste condizioni ricordiamo: la capacità, da parte della banca, di ricondurre i dati interni sulle perdite a linee di business e/o tipologie di evento specifiche; la capacità di dimostrare che la misura di rischio risultante dal modello sia stimata su un periodo di un anno e ad un livello di confidenza molto alto (99.9%); la sussistenza di un'unità indipendente di monitoraggio e controllo del rischio; l'attivo coinvolgimento del CdA e del *senior management* nella supervisione del processo di gestione del rischio operativo.

Nel disegno regolamentare di Basilea II, l'idea è dunque quella di incoraggiare le banche di maggior significatività e importanza a sviluppare il proprio modello interno per produrre dei risultati più *risk-sensitive* e che ne catturino meglio la struttura complessa e variegata rispetto a quelli offerti dagli approcci standardizzati, al contrario ritenuti più appropriati per banche dalla struttura più semplice. Dal punto di vista delle banche, infatti, la scelta di sviluppare e implementare un modello interno rappresenta sicuramente un investimento oneroso, la cui giustificazione risiede principalmente nel beneficio associato ad un requisito di capitale ridotto rispetto a quello che verrebbe calcolato attraverso i metodi standardizzati sin qui descritti.

Nonostante il BCBS non abbia predisposto modelli specifici, esso ha proposto comunque tre possibili approcci (non obbligatori) particolarmente rappresentativi degli AMAs: l'*internal measurement approach*, lo *scorecard approach*, e il *loss distribution approach.* Dei primi due metodi è fornita una concisa descrizione nel prosieguo del paragrafo, mentre al LDA è dedicata una spiegazione più estensiva in quanto, come anticipato, costituisce il metodo scelto per l'analisi contenuta in questo elaborato.

Internal Measurement Approach (IMA). Per utilizzare l'approccio IMA, una banca deve innanzitutto classificare le proprie attività in un dato numero j di linee di business (ad esempio

[43] Per l'elencazione di tutte le condizioni (generali, qualitative e quantitative): BCBS, "Calculation of RWA for operational risk: Advanced Measurement Approaches", Bank for International Settlements (2019)

otto come avviene per l'applicazione del TSA) e una gamma di k "tipologie di eventi" che possano dar luogo a perdite operative (ad esempio sette come indicato dal Comitato); la combinazione tra *event types/business lines* dà luogo ad una matrice, che nell'esempio menzionato è costituita da 56 "celle". Per ognuna di queste celle è poi calcolato lo specifico requisito di capitale come prodotto di quattro parametri:

$$RP_{IMA} = \sum_{j=1}^{8} \sum_{k=1}^{7} MI_{jk} \times PE_{jk} \times LGE_{jk} \times \gamma_{jk} \qquad (2.7)$$

dove:

- EI (*exposure indicator*) è l'indicatore di esposizione di pertinenza alla singola cella, cioè un indicatore specificato dall'autorità di vigilanza che costituisce una *proxy* per l'esposizione al rischio operativo. Un esempio d'indicatore può essere il margine d'intermediazione lordo;
- PE (*probability of event*) rappresenta la probabilità che il singolo evento si verifichi;
- LGE (*loss given the event*) indica l'entità della perdita correlata al verificarsi del singolo evento[44];
- Il parametro γ invece varia per ogni cella e rappresenta un fattore di "scala" rappresentativo della perdita inattesa (*unexpected loss, UL*); esso è fornito dall'autorità di vigilanza che lo calcola basandosi su dati settoriali.

Il prodotto tra i primi tre fattori (MI×PE×LGE) costituisce la perdita attesa (*expected loss, EL*) per ogni singola combinazione; moltiplicando poi la EL di ogni combinazione per il relativo fattore γ si ottiene la UL. Infine, sommando i risultati ottenuti per ogni cella della matrice si otterrà il requisito patrimoniale totale.

Questo approccio si basa su due importanti assunzioni, che tuttavia ne rappresentano anche il limite principale: la prima è che il metodo assume una perfetta correlazione tra ogni combinazione linea di business/tipologia di evento; la seconda assunzione è quella di relazione lineare tra la perdita attesa e quella inattesa, esplicitata dal parametro γ che è fisso e che potrebbe perciò non rappresentare adeguatamente l'effettivo profilo di rischio della banca.

Scorecard Approach. A differenza dell'IMA e, come vedremo, del LDA che sono metodi quantitativi, lo *Scorecard Approach* è una metodologia prettamente qualitativa. Essa consiste nel calcolo iniziale del requisito patrimoniale attraverso un metodo a scelta (ad esempio il BIA o il TSA), che viene poi attribuito alle singole linee di business sulla base di valutazioni

[44] Sia PE che LGE devono essere stimati dalla banca partendo dai propri dati interni.

(*scorecards*) dei profili di rischio e dei sistemi di controllo posti a presidio. Le valutazioni utilizzano degli indicatori rappresentativi dei particolari *event type* che si possono verificare in ogni linea di business, e vengono poi completate anche dal personale coinvolto nelle diverse aree di business ad intervalli regolari e infine sottoposte a revisione e validazione da parte di una *risk unit* centrale. Attraverso la *scorecard* viene poi attribuito un punteggio, diverso per ogni linea, attraverso cui la iniziale quota attribuita alla linea viene successivamente "ricalibrata". Si tratta dunque di un metodo che aggiunge un elemento *forward-looking* alla valutazione del rischio operativo. La scelta degli indicatori *proxy* è tipicamente affidata al giudizio degli esperti di settore.

2.3.4 Il cambio di rotta di Basilea III: *Standardised Measurement Approach*

Sebbene la riforma del *framework* di Basilea II iniziata nel 2010 (c.d. Basilea III) non aveva in principio interessato le sinora descritte metodologie di calcolo del requisito patrimoniale relativo al rischio operativo, un cambio di rotta si è avuto con la finalizzazione della riforma nel 2017. Nel corso dei lavori, infatti, il Comitato ha preso atto che i metodi previsti nel precedente accordo hanno fallito alla prova della grande crisi finanziaria del 2008: i requisiti patrimoniali calcolati si sono rivelati spesso insufficienti, e le metodologie di calcolo AMA si sono dimostrate troppo complesse e/o inconsistenti per giustificarne la riconferma. Da questa presa d'atto nasce dunque l'intenzione di una semplificazione totale del *framework*, con il ritiro totale degli approcci fin qui descritti in favore di un unico approccio standardizzato per tutti gli istituti finanziari; come si legge in un documento consultivo pubblicato dal Comitato nel 2016:

> *«L'esperienza di supervisione relativa agli AMA ha avuto risultati contrastanti. La difficoltà intrinseca di tali metodi e la mancanza di comparabilità dovuta ad una gamma di approcci interni vasta e variegata ha esacerbato la variabilità nel calcolo delle attività ponderate per il rischio ed eroso la fiducia nei coefficienti di capitale ponderati per il rischio.»*

Il nuovo approccio proposto dal Comitato si basa dunque sulla convinzione che la combinazione di una semplice e standardizzata misura di rischio operativo con i dati sulle perdite specifici delle singole banche sia idonea a garantire una misura sufficientemente *risk sensitive* che sia anche comparabile e di semplice implementazione[45]. Questa nuova metodologia, denominata *Standardised Measurement Approach* (SMA) combina dunque due elementi:

[45] BCBS (2016), "Standardised Measurement Approach for Operational Risk", Consultative Document, Bank for International Settlements.

$$RP_{SMA} = BIC \times ILM \qquad (2.8)$$

Il primo elemento, denominato *Business Indicator Component* (BIC) è ottenuto moltiplicando un indicatore *proxy* di bilancio, denominato *Business Indicator* (BI), per coefficienti marginali costanti e crescenti (che denomineremo μ) determinati in via regolamentare in base alla grandezza del BI. Quest'ultimo è a sua volta ottenuto combinando i valori medi degli ultimi tre anni di tre elementi: 1) gli interessi e dividendi (IDC – *Interest & Dividend Component*); 2) proventi da servizi, come ad esempio le commissioni (SC – *Service Component*); 3) componente finanziaria (FC), ossia il valore assoluto della somma tra profitti e perdite sul *Trading Book* e sul *Banking Book*. Per quanto riguarda i coefficienti marginali μ, come detto, questi sono applicati a scaglioni ('*buckets*')[46] in base al valore del BI; i valori di riferimento sono riassunti in tabella 3. In formule:

$$BIC = (IDC + SC + FC) \times \mu \qquad (2.9)$$

I coefficienti devono essere applicati in maniera progressiva, per cui ad esempio per un BI di €35 mld il calcolo da effettuare sarà il seguente:

$$BIC = (1 \times 12\%) + (30 \times 15\%) + (4 \times 18\%)$$

Bucket	BI range (in mld €)	Coefficiente Marginale μ
1	≤ 1	12%
2	1 ≤ BI ≤ 30	15%
3	> 30	18%

Tabella 3: Coefficienti Marginali per il calcolo del BIC. Fonte: BIS (2016), "Standardised Measurement Approach for operational risk"

Il secondo elemento del calcolo è denominato *Internal Loss Multiplier* (ILM) ed è un fattore di scala che serve ad "aggiustare" il requisito di capitale sulla base dell'esperienza della banca in fatto di perdite operative. Si tratta dunque di una componente *risk sensitive* in quanto basata sui dati interni agli istituti finanziari ed è calcolata in base alla seguente formula:

$$ILM = \ln\left(\exp(1) - 1 + \left(\frac{LC}{BIC}\right)^{0.8}\right) \qquad (2.10)$$

In questa formula figurano sia il BIC che una componente di perdita (*loss component* – LC), pari a 15 volte la perdita operativa media annuale calcolata sulla base dei precedenti 10

[46] Originariamente, i *buckets* previsti erano cinque.

anni[47]. Il rapporto fra queste due componenti darà luogo a un ILM maggiore, minore o pari a uno se, rispettivamente:

- LC > BIC, che significa anche che le perdite operative della banca sono alte in relazione al BIC e dunque la banca è tenuta a detenere maggior capitale;
- LC < BIC, cioè le perdite operative sono relativamente basse e dunque la banca è tenuta a detenere minor capitale;
- LC = BIC, nel qual caso la componente ILM diventa ininfluente ai fini del calcolo del requisito patrimoniale.

È importante sottolineare infine che la riforma lascia spazio alla discrezionalità delle autorità nazionali riguardo all'ILM, che può essere settato pari a 1 per tutte le banche in modo da annullare il contributo delle perdite interne al calcolo del requisito.

L'entrata in vigore definitiva del nuovo set di regole è prevista per il 1° gennaio 2023, spostata in avanti di un anno rispetto alle previsioni iniziali a causa della pandemia da Covid-19, sebbene molti degli standard previsti da Basilea III siano già attualmente in vigore in alcune giurisdizioni.

2.4 Il Loss Distribution Approach (LDA)

Il terzo ed ultimo metodo proposto dal Comitato di Basilea come "rappresentante" degli approcci AMA è il Loss Distribution Approach, una metodologia derivata dalle scienze attuariali che si è nel tempo affermata come lo standard di riferimento nell'ambito della modellizzazione del rischio operativo. Come si legge in un documento consultivo a supporto dell'accordo di Basilea II[48], una banca che utilizzi il LDA deve stimare una funzione di distribuzione di probabilità per la severità e per la frequenza di ogni combinazione (o "cella") tra *business line/event type* su un orizzonte di un anno, combinando poi le due distribuzioni per ottenere una funzione di distribuzione cumulata per le perdite operative per ogni combinazione. In formule[49],

$$Z_j = \sum_{i=1}^{N_j} X_i^{(j)} = X_1^{(j)} + X_2^{(j)} + \cdots + X_{N_j}^{(j)} \qquad (2.11)$$

[47] Il calcolo della LC deve essere basato su dati di «elevata qualità» che devono rispettare alcuni criteri stabiliti dal Comitato; per maggiori informazioni, https://www.bis.org/basel_framework/chapter/OPE/25.htm?inforce=20230101&published=20200605

[48] BCBS, "Operational Risk" Supporting Document to the New Basel Capital Accord [Appendice 6], Bank for International Settlements (2001)

[49] Le equazioni (2.11) e (2.12) sono riprese e adattate da Shevchenko (2010).

dove,

- Le $X_i^{(j)}$ sono le variabili casuali che rappresentano la severità delle perdite per la j-esima cella;
- N_j è la variabile casuale che rappresenta la frequenza delle perdite per la j-esima cella;
- Z_j è quindi la variabile casuale che risulta dalla combinazione delle precedenti due, ossia la funzione di distribuzione cumulata delle perdite operative annuali per la j-esima combinazione

La (2.11) assume implicitamente che le v.c. X (che rappresentano l'impatto dei singoli eventi di perdita) siano indipendenti e identicamente distribuite tra loro, e che la frequenza N_j e la *severity* X_i siano indipendenti.

Generalmente parlando, è improbabile che si ottenga una forma analitica nota e trattabile per la distribuzione cumulata delle perdite Z_j, per cui è necessario ricorrere ad algoritmi numerici, come il metodo MonteCarlo, per ottenerne un'approssimazione (la cui bontà dipende, ovviamente, dal metodo applicato). Aggregando le distribuzioni cumulate delle varie celle, si ottiene la distribuzione cumulata delle perdite operative totali per la banca:

$$Z = \sum_{j=1}^{J} Z_j \qquad (2.12)$$

Una volta ottenuta la distribuzione cumulata delle perdite, il passo successivo consiste nel calcolo del VaR; riprendendo la (2.1)[50]:

$$VaR_\alpha(Z) = \inf\{z \in \mathbb{R} : P(Z > z) \leq (1-\alpha)\} = \inf\{z \in \mathbb{R} : F_Z(z) \geq \alpha\} \qquad (2.13)$$

Il requisito di capitale equivarrà dunque al VaR calcolato ad un livello di confidenza del 99,9% sulla distribuzione cumulata delle perdite operative totali. È importante notare che il procedimento sin qui descritto implica una semplice aggregazione dei profili di rischio delle varie celle, tralasciando completamente eventuali correlazioni che vi possono essere tra queste ultime e che potrebbero potenzialmente portare ad un risultato diverso e più accurato qualora venissero considerate. Tuttavia, la stima di tali correlazioni potrebbe rivelarsi troppo incerta o difficoltosa, motivo per il quale il BCBS suggerisce di ignorarle[51].

[50] Il livello di confidenza α è fissato al 99.9%.

[51] Una banca potrebbe comunque essere autorizzata dall'Autorità di supervisione a includere nella determinazione del requisito di capitale le correlazioni tra le varie *risk cell*, calcolate internamente, a condizione che riesca a dimostrare che i sistemi impiegati nella determinazione di tali correlazioni siano affidabili e che tengano in considerazione l'incertezza connaturata alle stime.

Come accennato in apertura, il *Loss Distribution Approach* si è nel tempo affermato come il *framework* di riferimento per affrontare i problemi legati alla valutazione e alla gestione del rischio operativo; si tratta comunque di un approccio non esente da limitazioni, e nel tentativo di superarle diverse soluzioni e metodi differenti sono stati proposti in letteratura.

Un primo problema è relativo alla ben nota scarsità di dati utilizzabili per la stima delle distribuzioni di frequenza e severità delle perdite, in quanto a differenza del rischio di mercato o di credito, la raccolta di dati per il rischio operativo è iniziata in tempi relativamente più recenti. Per superare questo importante ostacolo, le soluzioni più diffuse consistono nell'utilizzo di distribuzioni parametriche a cui adattare i dati, e l'aggregazione di dati interni e di dati esterni (dunque provenienti da fonti diverse) per espandere i dataset disponibili. Nel primo caso si procede dunque a cercare di adattare i dati disponibili ad alcune distribuzioni di probabilità note, stimandone i parametri e la relativa incertezza di stima, verificando poi la bontà di adattamento del modello. L'aggregazione di dati interni e di dati esterni è invece un requisito richiesto in Basilea II per la qualificazione dei modelli AMA e si tratta di un processo delicato; una soluzione al problema è proposta in Shevchenko (2010).

Tuttavia, è opportuno rimarcare che anche l'utilizzo e l'aggregazione di dati provenienti da fonti diverse non riesce a risolvere tutte le problematiche connesse alla scarsità di dati. Come sottolineato in Frachot, Georges and Roncalli (2001) il modo stesso in cui questi vengono raccolti è fonte di problemi: è probabile, infatti, che solo le perdite più significative (quindi al di sopra di una certa soglia) vengano registrate, portando ad una sovrastima della severità delle perdite. Questo problema è noto in letteratura come *truncation bias*.

Un secondo problema, strettamente collegato al primo, riguarda invece la natura dei dati sulle perdite operative. È infatti noto che, molto più di altre tipologie di rischio, il rischio operativo si caratterizza per perdite di notevole entità che però occorrono con frequenza molto bassa, contrapposte a perdite di severità contenuta che occorrono con frequenza più alta. Ciò comporta inevitabilmente una difficoltà ulteriore nella costruzione di modelli adeguati, in quanto i dati riguardanti le perdite più severe potrebbero essere pochi o addirittura assenti nei dataset a disposizione. Per questo motivo, nella modellizzazione della *severity* delle perdite, una delle soluzioni più comunemente adottate (oltre all'utilizzo di dati interni ed esterni) comporta l'utilizzo di una distribuzione giunta (*spliced distribution*) ove il "corpo" della distribuzione, relativo alle perdite più frequenti e per cui sono quindi disponibili più dati, è modellato secondo una distribuzione comune (ad esempio la Weibull, la Log-Normale o anche la distribuzione empirica), mentre la coda della distribuzione, relativa alle perdite più rare ma di maggior impatto è tipicamente caratterizzata attraverso tecniche della *Extreme Value Theory*.

Definizione 2. Una v.c. X ha una distribuzione giunta (o *spliced distribution*) se la sua funzione di densità è data da

$$f_X(x) = \begin{cases} w_1 f_1(x), & x_0 \le x < x_1 \\ w_2 f_2(x), & x_1 \le x < x_2 \\ \dots & \\ w_k f_k(x), & x_{k-1} \le x < x_k \end{cases} \qquad (2.14)$$

dove le $w_j > 0$ rappresentano i pesi attribuiti alle diverse distribuzioni; ne segue ovviamente che

$$\sum_j w_j = 1$$

Nel resto del capitolo saranno dunque esaminate alcune delle distribuzioni tipicamente utilizzate in letteratura per la modellizzazione della severità e della frequenza delle perdite nell'ambito del rischio operativo e del *cyber risk*; in chiusura sono poi forniti alcuni elementi utili di EVT.

2.4.1 Distribuzioni tipiche di frequenza

Il primo passo per l'implementazione del LDA consiste nella stima della distribuzione di frequenza delle perdite. Come accennato nel paragrafo precedente, l'approccio parametrico è quello preferenziale per risolvere i problemi legati alla scarsa qualità o disponibilità dei dati; è opportuno dunque presentare le distribuzioni parametriche più utilizzate per la stima della frequenza delle perdite.

La **distribuzione di Poisson** è senza dubbio la scelta principale e più diffusa nella letteratura sul rischio operativo[52] e in quella più recente sul *cyber risk*[53]. Si tratta di una distribuzione di probabilità discreta, utilizzata per trovare la probabilità del verificarsi di un certo numero *k* di eventi in un dato intervallo temporale; indicando con λ il numero medio di eventi in questo intervallo, allora la probabilità cercata è data da

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}, \qquad k = 0,1,2,\dots \qquad (2.15)$$

La distribuzione di Poisson presenta alcuni notevoli vantaggi. Il primo consiste nella dipendenza da un unico parametro, λ (chiamato anche *intensity rate*), che ne rappresenta sia il valore atteso che la varianza, a favore dunque di una maggiore semplicità:

[52] Ad esempio, Frachot, Georges and Roncalli (2001) o Shevchenko (2010)
[53] Bouveret (2018) o Eling and Wirfs (2019)

$$E(X) = \lambda, \;\; Var(X) = \lambda$$

Un altro importante vantaggio è dato dal fatto che, supponendo che X e Y siano due v.c. indipendenti e descritte entrambe da una Poisson con parametri λ_X e λ_Y, allora la distribuzione di X + Y è a sua volta una Poisson con parametro $\lambda_X + \lambda_Y$; si tratta di una proprietà utile quando, ad esempio, si voglia valutare la frequenza delle perdite da due o più *business line*. Infine, come evidenziato da Panjer (2006)[54], se la frequenza delle perdite si distribuisce secondo una Poisson, e le perdite possono essere classificate in un numero *n* di categorie, seguirà che anche le perdite in ogni categoria seguono delle Poisson con parametri λ diversi. Questa proprietà è di particolare utilità nello studio dei rischi, in quanto ad esempio spesso le perdite vengono classificate in base ad una determinata soglia *u*: grazie a questa proprietà, è possibile dunque studiare sia la frequenza delle perdite al di sopra che al di sotto della soglia utilizzando la distribuzione di Poisson ma stimandone parametri diversi.

Nonostante la semplicità, l'utilizzo di tale distribuzione richiede comunque alcune cautele. Un importante limite discusso in Frachot, Moudoulaud and Roncalli (2003) segue dal *truncation bias* di cui si è discusso nel paragrafo precedente; si è accennato infatti al fatto che, generalmente, i dati vengono registrati solo per perdite al di sopra di una certa soglia di "significatività" implicita, mentre le perdite al di sotto di tale soglia non vengono registrate. Ciò comporta, logicamente, che anche con un campione sufficientemente ampio e di "buona qualità" vi sarà una sottostima della frequenza delle perdite con conseguente sottostima del requisito patrimoniale. La soluzione introdotta dagli stessi autori consiste nella correzione del parametro λ attraverso una semplice osservazione: il rapporto tra il numero di eventi osservato e il reale numero di eventi è esattamente pari alla probabilità che un evento porti ad una perdita superiore alla soglia implicita di cui sopra. In formule,

$$\lambda_{reale} = \frac{\lambda_{campionario}}{P(L > h)} \qquad (2.16)$$

dove L è la perdita e h è la soglia implicita di registrazione. Ciò implica anche che la stima della distribuzione di frequenza delle perdite debba avvenire in un momento successivo alla stima della distribuzione di severità delle stesse, in quanto solo così sarà possibile derivare il denominatore della (2.16).

Un secondo problema riguarda poi la natura del parametro λ, che è assunto essere costante. Si tratta evidentemente di un'assunzione poco realistica, in quanto è probabile che

[54] Panjer, H. (2006) "*Operational Risk*" (Capitolo 5, Teorema 5.2)

esso vari nel tempo o che abbia un comportamento randomico. Per rimuovere questa assunzione poco realistica, una possibile soluzione consiste nell'assumere che λ segua un processo stocastico, e che dunque si evolva nel tempo secondo una funzione matematica λ(*t*) (processo di Poisson non-omogeneo con intensità stocastica); oppure si può assumere che λ segua una distribuzione di probabilità a sua volta, ottenendo una mistura di distribuzioni (*mixture distribution*).

Un caso speciale di mistura di distribuzioni è la distribuzione Binomiale Negativa (BN)[55]; infatti, se si assume che il parametro λ della Poisson segua a sua volta una distribuzione Gamma, si ottiene proprio tale distribuzione. A differenza della Poisson, la Binomiale Negativa è una distribuzione tipicamente descritta da due parametri e che dunque ammette una maggiore flessibilità nella forma. Intuitivamente, essa rappresenta il numero di "fallimenti" che incorrono in una serie di prove di Bernoulli prima di ottenere un determinato numero di successi. La funzione di probabilità della BN è la seguente[56]:

$$f(x) = \frac{\Gamma(x+n)}{x!\,\Gamma(n)} p^n (1-p)^x \qquad (2.17)$$

$$x \in [0, \infty[\,, \qquad n > 0, \qquad 0 < p \leq 1$$

In questa funzione, il parametro n è un parametro di "dispersione" (più avanti definio anche *size*), mentre p rappresenta la probabilità di successo in ogni prova di Bernoulli. Questa funzione può anche essere parametrizzata in termini della media μ, sfruttando il fatto che

$$p = \frac{n}{n+\mu}$$

Nel Capitolo 3 saranno fornite le stime dei parametri n e μ per i dati su cui sarà svolta l'analisi.

2.4.2 Distribuzioni tipiche di *severity*

La stima della distribuzione di probabilità della dimensione delle perdite è un passaggio un po' più delicato nell'applicazione della metodologia. Infatti, mentre nella stima della frequenza è ragionevole e non troppo costoso – in termini di accuratezza delle stime – fare delle assunzioni semplificatrici (qual è l'adozione della Poisson, ad esempio), lo stesso non vale per questa fase del processo; come affermato da Frachot, Moudoulaud and Roncalli (2003) non è possibile applicare le classiche tecniche di analisi statistica senza deteriorare la qualità della

[55] Chernobai, Rachev and Fabozzi (2012).
[56] Questa è la formulazione che si ottiene dall'interpretazione della Binomiale Negativa come mistura di una Poisson e di una distribuzione Gamma; altre formulazioni sono comunque possibili.

stima, in quanto i dati che si vanno a trattare sono generalmente distorti o insufficienti e si viene dunque a configurare un *trade-off* tra accuratezza e semplicità. Ciò comporta, tra le altre cose, la virtuale impossibilità di utilizzare la distribuzione empirica delle perdite[57] ai fini della stima e che si debba dunque ricorrere a distribuzioni parametriche da adattare ai dati disponibili.

In questo paragrafo è presentata dunque una panoramica delle distribuzioni di probabilità più popolari derivata da Chernobai, Rachev and Fabozzi (2012) e Cruz, Peters and Shevchenko (2015). È necessario comunque fare alcune precisazioni di carattere generale.

Innanzitutto, il focus di quest'analisi riguarda modelli utilizzati per caratterizzare la severità delle perdite, e appare dunque poco sensato l'utilizzo di distribuzioni che possano assumere valori sia positivi che negativi (come la distribuzione Normale): una "perdita negativa" non è altro che un profitto. Per questo motivo, dunque, saranno presentate solamente distribuzioni definite su un supporto non-negativo, in quanto non vi è interesse nella caratterizzazione dei profitti. Un altro fatto da tenere a mente è che è decisamente infrequente l'utilizzo di una singola distribuzione parametrica per la modellizzazione delle perdite; come accennato in precedenza, l'approccio più popolare è quello di utilizzare distribuzioni diverse per caratterizzare il "corpo" e la "coda", ossia le perdite più frequenti e quelle più rare, della distribuzione di *severity*.

Distribuzione Esponenziale – Exp (θ). Si tratta di una distribuzione descritta da un singolo parametro θ con funzione di probabilità e funzione di distribuzione, rispettivamente,

$$f(x) = \theta e^{-\theta x}, \qquad F(x) = 1 - e^{-\theta x}$$

con x > 0.

I momenti semplici della distribuzione possono essere calcolati in base alla formula

$$E(X^k) = \frac{k!}{\theta^k}$$

E dunque valore atteso e varianza saranno pari, rispettivamente, a

$$E(X) = \frac{1}{\theta} \qquad E(X^2) = Var(X) = \frac{1}{\theta^2}$$

[57] L'utilizzo della distribuzione empirica, infatti, presuppone due importanti assunzioni: un dataset sufficientemente ampio e completo; l'assunzione che tutte le perdite passate possano verificarsi di nuovo, ma che perdite di diversa entità (non contenute nel dataset) non possano verificarsi. Quest'ultima assunzione, in particolare, è molto forte e non si presta bene allo studio del rischio operativo. Fonte: Chernobai, Rachev and Fabozzi (2012).

La distribuzione esponenziale attribuisce una probabilità molto bassa, quasi pari a 0, a perdite di particolare severità in quanto presenta una coda destra che decade esponenzialmente. Per questo motivo non è particolarmente adatta per la caratterizzazione del rischio operativo.

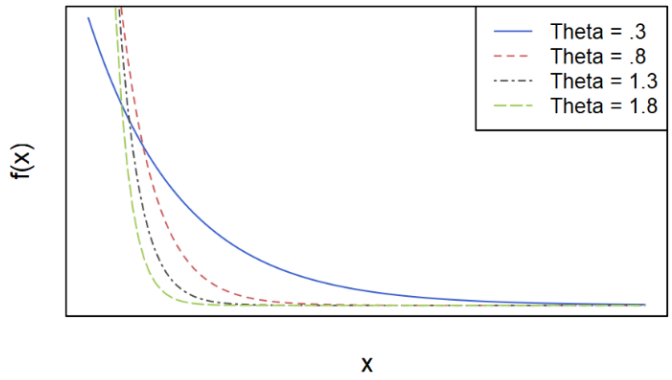

Figura 3: Densità della Distribuzione Esponenziale al variare del parametro θ

Distribuzione Log-Normale – LogNormal (μ, σ²). La Log-Normale è una distribuzione caratterizzata da code pesanti (*heavy-tailed*)[58] e descritta da due parametri μ e σ^2, rispettivamente parametro di posizione (*location*) e di scala. La funzione di probabilità e la funzione di distribuzione sono, rispettivamente,

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma x} e^{-\frac{(\log(x)-\mu)^2}{2\sigma^2}}, \qquad F(x) = \Phi\left(\frac{\log(x)-\mu}{\sigma}\right)$$

con $x > 0$ e dove $\Phi(x)$ rappresenta la funzione di distribuzione di una Normale standard.
I momenti semplici della distribuzione log-normale possono essere calcolati come

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

Valore atteso e varianza saranno dunque pari a

$$E(X) = e^{\mu + \frac{\sigma^2}{2}} \qquad E(X^2) = Var(X) = \left(e^{\sigma^2} - 1\right)e^{2\mu+\sigma^2}$$

Tale distribuzione può essere ottenuta mediante trasformazione di una Gaussiana; infatti, se $Y \sim N(\mu, \sigma^2)$ allora si avrà che $e^Y = X \sim LogNormal(\mu, \sigma^2)$.

[58] Per *heavy-tailed distribution* si intende comunemente una distribuzione le cui code non sono limitate esponenzialmente; in altre parole, si tratta di una distribuzione le cui code sono più pesanti di quelle della distribuzione esponenziale.

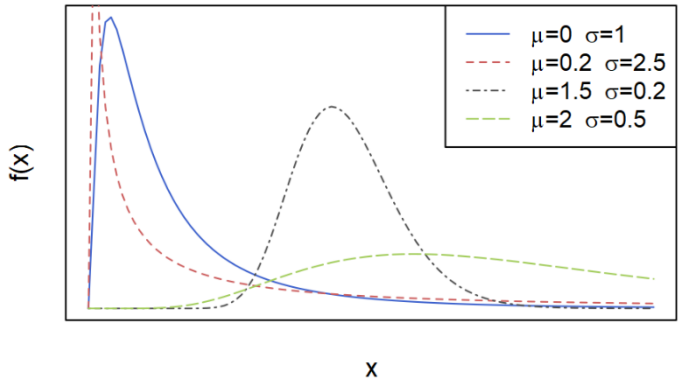

Figura 4: Densità della distribuzione Log-Normale al variare dei parametri μ e σ.

Distribuzione di Weibull – Weibull (α, β). Una variabile aleatoria X ha una distribuzione di Weibull se la sua funzione di densità di probabilità e funzione di distribuzione sono, rispettivamente

$$f(x) = \alpha\beta x^{\alpha-1}e^{-\beta x^{\alpha}} \qquad F(x) = 1 - e^{-\beta x^{\alpha}}$$

con x > 0, α > 0 (parametro di forma), β > 0 (parametro di scala).

La Weibull è una generalizzazione della distribuzione esponenziale, a cui si riduce nel caso in cui α = 1; nel caso in cui $\alpha \in (0,1)$, essa è caratterizzata da code pesanti. Valore atteso e varianza sono rispettivamente

$$E(X) = \beta^{-\frac{1}{\alpha}}\Gamma\left(1+\frac{1}{\alpha}\right) \qquad E(X^2) = Var(X) = \beta^{-\frac{2}{\alpha}}\left(\Gamma\left(1+\frac{2}{\alpha}\right) - \Gamma^2\left(1+\frac{1}{\alpha}\right)\right)$$

dove Γ è una funzione gamma.

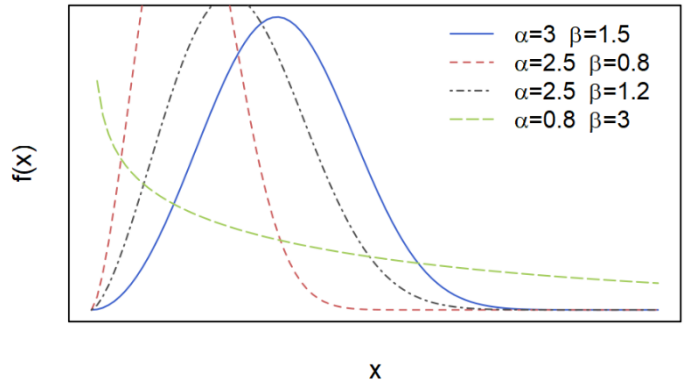

Figura 5: Densità della distribuzione di Weibull al variare dei parametri α e β

Distribuzione Paretiana a due parametri[59] – Pareto (α, β). La distribuzione Paretiana è descritta dalle seguenti funzioni di densità e di distribuzione

$$f(x) = \frac{\alpha\beta^{\alpha}}{x^{\alpha+1}} \qquad F(x) = 1 - \left(\frac{\beta}{x}\right)^{\alpha}$$

Con $0 < \beta < x < \infty$ e $\alpha > 0$, rispettivamente parametro di scala e di forma.

I momenti ordinari possono essere stimati dalla funzione

$$E(X^k) = \frac{\alpha\beta^k}{\alpha - k}$$

e quindi valore atteso e varianza sono espressi nel modo che segue

$$E(X) = \frac{\alpha\beta}{\alpha - 1} \qquad E(X^2) = \frac{\alpha\beta^2}{\alpha - 2}$$

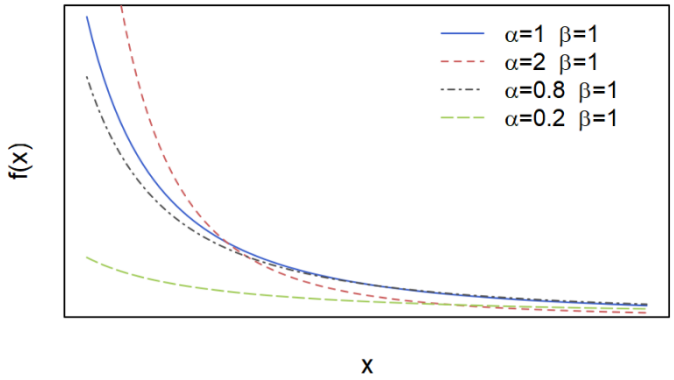

Figura 6: Densità della distribuzione di Pareto al variare del parametro α.

Si tratta di una distribuzione caratterizzata da una coda molto pesante, il cui comportamento è determinato dal parametro α in modo che più esso è vicino a zero, più spessa sarà la coda (vedi Figura 6); per valori di $\alpha \leq 1$, come è possibile notare anche dalle rispettive formule, media e varianza diventano infinite ammettendo dunque la possibilità del verificarsi di perdite di entità potenzialmente illimitata, una proprietà non particolarmente attrattiva nello studio dei rischi. Attraverso una riparametrizzazione della Paretiana è possibile ottenere la distribuzione di Pareto generalizzata (*Generalized Pareto Distribution – GPD*), spesso utilizzata per modellare la coda della distribuzione di *severity* attraverso l'approccio POT (*Peaks-over-Treshold*) della EVT.

[59] È esplicitato il riferimento ai due parametri in quanto esiste anche una distribuzione Paretiana con un singolo parametro, ottenibile attraverso trasformazione di una Esponenziale.

2.5 Lo studio delle code: *Extreme Value Theory*

La *Extreme Value Theory* (EVT) è una branca della statistica che si occupa, come intuibile dal nome, dello studio dei valori estremi delle distribuzioni di probabilità – ossia quei valori che si allontanano fortemente dalla porzione centrale delle stesse (il "corpo"). Le tecniche della EVT sono spesso utilizzate nello studio dei rischi, e in particolare nello studio del rischio operativo, in quanto conferiscono la possibilità di valutare le code di una distribuzione anche in presenza di dati limitati. In altre parole, la EVT si occupa dello studio degli eventi più rari, caratterizzati da una bassa probabilità di verificarsi, ma che possono avere conseguenze potenzialmente catastrofiche. Per questo motivo, si tratta di una teoria particolarmente attrattiva per coloro che operano nel campo della gestione dei rischi, spesso chiamati a tenere conto di questo tipo di eventi nelle loro valutazioni. La EVT si fonda essenzialmente due tipi di modelli, o approcci, fondamentali: l'approccio *Block Maxima* (BM) e il relativamente più moderno approccio *Peaks-over-Treshold* (POT) cui si è accennato alla fine del precedente paragrafo, che rappresenta anche quello più diffuso nello studio dei rischi operativi in quanto consente di ottimizzare l'uso di dataset limitati. Di seguito è dunque fornita una breve descrizione dei due approcci, ove per una trattazione più completa si rimanda a McNeill (1999) o Peters and Shevchenko (2015).

2.5.1 Il metodo *Block Maxima.*

Nell'approccio *Block Maxima* si procede essenzialmente alla suddivisione dei dati in "blocchi" temporali di uguale ampiezza, e da ogni blocco si estrae la singola osservazione di maggior entità (ossia il massimo); si dimostra che, per valori sufficientemente alti, la distribuzione di tali massimi "normalizzati" estratti dai blocchi converge alla *Generalized Extreme Value Distribution* (GEV), la cui funzione di distribuzione è di seguito definita:

$$F_{GEV}(x) = \begin{cases} \exp\left(-(1+\xi x)^{-1/\xi}\right), & \xi \neq 0 \\ \exp\left(-e^{-x}\right), & \xi = 0 \end{cases}$$

Tale risultato, che va sotto il nome di teorema di Fisher-Tippett, ricopre un'importanza cruciale nell'ambito dell'EVT, similarmente al Teorema del Limite Centrale nella statistica classica. In base al teorema, dunque, la GEV è l'unica distribuzione limite possibile per i *block maxima* (normalizzati); essa può assumere diverse conformazioni, a seconda del valore di ξ:

- per valori di $\xi > 0$ si ottiene una distribuzione di Fréchet;
- per $\xi = 0$, si ottiene una distribuzione di Gumbel;
- con $\xi < 0$, si ha una distribuzione di Weibull.

2.5.2 Il metodo *Peaks-over-Threshold*

L'approccio POT consiste nello studio dei valori che superano una determinata soglia *u*, fissata ad un livello sufficientemente alto. Si assuma che X sia una variabile casuale rappresentativa delle perdite operative e che F sia la relativa funzione di distribuzione; con F_u si indica dunque la funzione di distribuzione delle perdite al di sopra della soglia scelta, definita come

$$F_u(y) = P(X - u \leq y | X > u) = \frac{F(u+y) - F(u)}{1 - F(u)} \qquad (2.18)$$

$F_u(y)$ rappresenta dunque la **distribuzione delle perdite in eccesso**, ossia la probabilità che una perdita superi la soglia *u* di un ammontare pari al massimo a *y*, posto che essa superi effettivamente la soglia; in altre parole, essa definisce la distribuzione di probabilità della coda destra. Uno dei risultati chiave della EVT, che va sotto il nome di Teorema di Pickands-Balkema-de Haan, consiste nel fatto che, per un'ampia gamma di distribuzioni[60], $F_u(y)$ converge ad una distribuzione di Pareto generalizzata (GPD) all'aumentare della soglia *u*.
In simboli:

$$\lim_{u \to x_0} \sup_{0 \leq x \leq x_o - u} \left| F_u(x) - G_{\xi,\beta}(x) \right| = 0$$

La funzione di distribuzione della GPD è definita come segue:

$$G_{\xi,\beta}(x) = \begin{cases} 1 - \left(1 + \xi \frac{x-\mu}{\beta}\right)^{-\frac{1}{\xi}}, & \xi \neq 0 \\ 1 - e^{-\frac{x-\mu}{\beta}}, & \xi = 0 \end{cases}$$

Dove,

- *x* rappresenta le osservazioni al di sopra della soglia *u*;
- $-\infty < \mu < +\infty$ è un parametro di posizione. Esso è spesso assunto pari a 0, motivo per il quale non compare affatto in alcune formulazioni[61].
- $\beta > 0$ è il parametro di scala;
- ξ è il parametro che determina la forma della distribuzione: per valori di $\xi > 0$ otteniamo una distribuzione *heavy-tailed* come riparametrizzazione di una ordinaria distribuzione di Pareto, già descritta nel paragrafo precedente; con $\xi = 0$ si ottiene una

[60] Sono incluse tutte le distribuzioni più comunemente utilizzate in statistica e scienze attuariali, tra cui ad esempio Normale, Log-Normale, Gamma, Esponenziale, F di Fisher, ecc. (McNeill 1999).
[61] Ad esempio, non compare nella formulazione data da McNeill (1999).

semplice distribuzione esponenziale; perer valori di $\xi < 0$ si ottiene una distribuzione nota come Paretiana di II tipo.

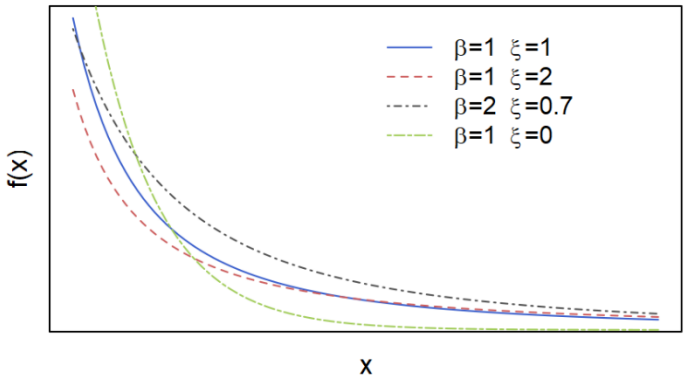

Figura 7: Densità della GPD al variare dei parametri di scala e forma; da notare che con $\xi = 0$ si ottiene esattamente il grafico di una distribuzione esponenziale.

Grazie al risultato appena illustrato è possibile dunque considerare la GPD come la scelta più ovvia per andare a modellare la distribuzione delle perdite in eccesso oltre una certa soglia. Tuttavia, come si sarà intuito, è proprio la scelta del valore soglia *u* a rappresentare uno dei passaggi chiave dell'applicazione del metodo; come evidenziato da McNeill (1999), bisogna essenzialmente trovare un compromesso tra la scelta di un valore sufficientemente alto, tale da permettere l'applicazione delle proprietà asintotiche appena descritte, e la scelta di un valore abbastanza basso da avere dati sufficienti per la stima dei parametri della distribuzione. Si tratta di un problema per il quale tuttavia non vi sono ancora soluzioni universalmente accettate nella pratica: un possibile approccio, frequentemente utilizzato, si basa sull'analisi visiva del grafico della *mean excess function* (funzione dell'eccesso medio), definita come la media di tutte le differenze tra i valori che superano la soglia *u* e *u*, per diversi valori della soglia.

$$MEF(u) = E(X - u | X > u)$$

Una volta scelta la soglia ottimale e stimati i parametri, il passo successivo consiste nella stima del VaR. McNeill (1999), supponendo che per le perdite al di sotto della soglia sia utilizzata la distribuzione empirica, fornisce la seguente formula per la stima del valore a rischio attraverso il metodo POT:

$$\widehat{VaR_\alpha}(X) = u + \frac{\hat{\beta}}{\hat{\xi}}\left(\left(\frac{n}{N_u}(1-\alpha)\right)^{-\hat{\xi}} - 1\right) \qquad (2.19)$$

dove con N_u è indicato il numero di osservazioni al di sopra della soglia *u*, mentre *n* rappresenta il numero totale di osservazioni. Per la derivazione completa della formula si rimanda all'articolo originale o a Moscadelli (2004).

2.6 *Goodness-of-Fit* e selezione del modello

La scelta delle distribuzioni che meglio rappresentano la frequenza e l'onerosità delle perdite costituisce, senza dubbio, la parte più delicata e complessa della costruzione di un modello di quantificazione dei rischi. Nel corso del capitolo si è fatto riferimento ad alcune semplici distribuzioni e modelli teorici che tipicamente utilizzati in questo processo, e innumerevoli altri sono stati proposti nella letteratura scientifica dedicata; del resto, appare chiaro che non è possibile ricavare un modello inequivocabilmente superiore agli altri, ma che piuttosto bisogna puntare a costruire un modello sufficientemente accurato sulla base degli strumenti e dei dati che si hanno a disposizione e attraverso un meccanismo che possiamo definire di *trial-and-error*; un processo di questo tipo passa, necessariamente, attraverso la conferma o la smentita delle ipotesi fatte a monte.

In questa fase è dunque utile fare riferimento ad alcune "tecniche" che consentano in qualche modo di valutare la bontà dell'analisi svolta e del modello teorico che ne risulta. È tuttavia necessario anche sottolineare che valutazioni di questo tipo non possono essere schematiche ed esatte in tutti i casi: determinate tecniche saranno più utili in alcuni casi piuttosto che in altri, a seconda della strategia analitica adottata e della domanda a cui si cerca di dare una risposta con la stessa.

Il primo passo nella conduzione di questo processo consiste nell'analisi grafica: attraverso l'osservazione di alcuni particolari grafici è infatti possibile iniziare a cogliere informazioni utili per il prosieguo dello studio, informazioni che poi dovranno essere confermate o smentite attraverso metodi numerici più formali; a dispetto di ciò, tale metodo rappresenta un ottimo punto di partenza in quasi ogni circostanza. Il primo dei grafici che generalmente si va ad osservare è quello della funzione di distribuzione cumulata empirica (ECDF), detta anche funzione di ripartizione empirica (figura 8).

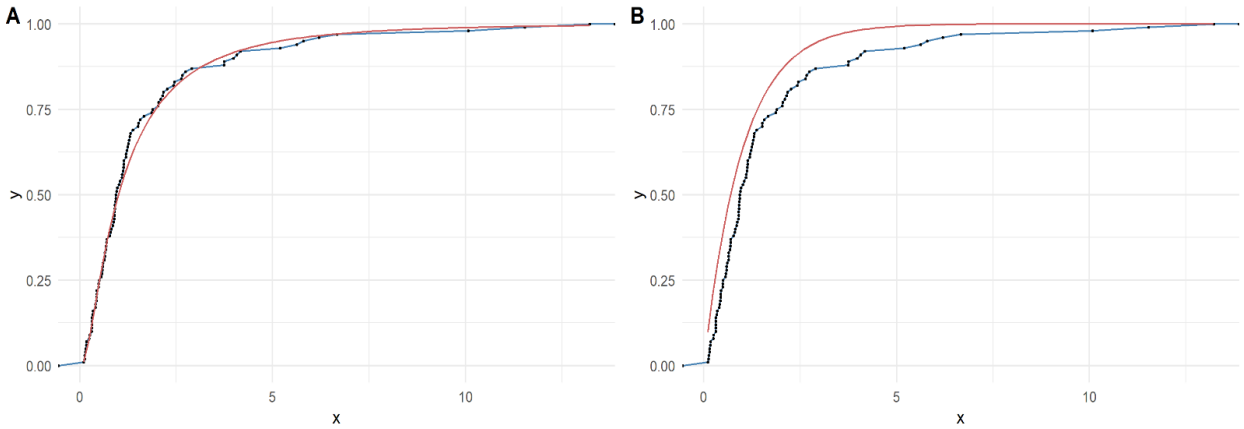

Figura 8. Nel pannello (A) è presente il confronto fra una funziona di distribuzione cumulata empirica confrontata con la funzione di distribuzione teorica di una distribuzione log-normale. Nel pannello (B) la stessa ECDF è confrontata con una distribuzione esponenziale. Fonte: elaborazione dell'autore.

Nel grafico (pannello A) è confrontata la ECDF (linea punteggiata) di alcuni dati generati casualmente da una distribuzione log-normale con la CDF teorica della stessa distribuzione (linea continua): come è possibile notare, il *fit* tra le due curve è quasi perfetto poiché provengono dalla stessa distribuzione; nel pannello B, invece, la stessa ECDF è confrontata con la CDF teorica di una distribuzione esponenziale con parametro $\theta = 1$: si può notare che non vi è praticamente sovrapposizione tra le due curve, e anzi quella empirica si trova costantemente al di sotto della teorica indicando dunque che i dati provengono da una distribuzione *heavy-tailed*, come appunto è la log-normale. In questo semplice scenario, dunque, l'utilizzo di un modello esponenziale porterebbe ad una costante sottostima delle probabilità.

Oltre al grafico della CDF, un'ulteriore possibilità consiste nel confrontare l'istogramma dei dati con la densità della distribuzione stimata; tuttavia, il ricorso a tale grafico può risultare più problematico e meno immediato in alcuni casi, in quanto potrebbero occorrere problemi nel raggruppamento dei dati per generare l'istogramma. Ad ogni modo, entrambi i metodi risultano particolarmente sensibili alla presenza di *outliers* che rischiano di distorcere le rappresentazioni grafiche.

Un altro tipo di rappresentazione tipicamente molto utile è il *Q-Q plot*; in un grafico di questo tipo, i quantili della distribuzione empirica vengono confrontati con i quantili della distribuzione teorica ipotizzata. Intuitivamente, il *fit* fra le due distribuzioni sarà tanto migliore quanto più vicini saranno i punti del grafico ad una retta inclinata a 45°.

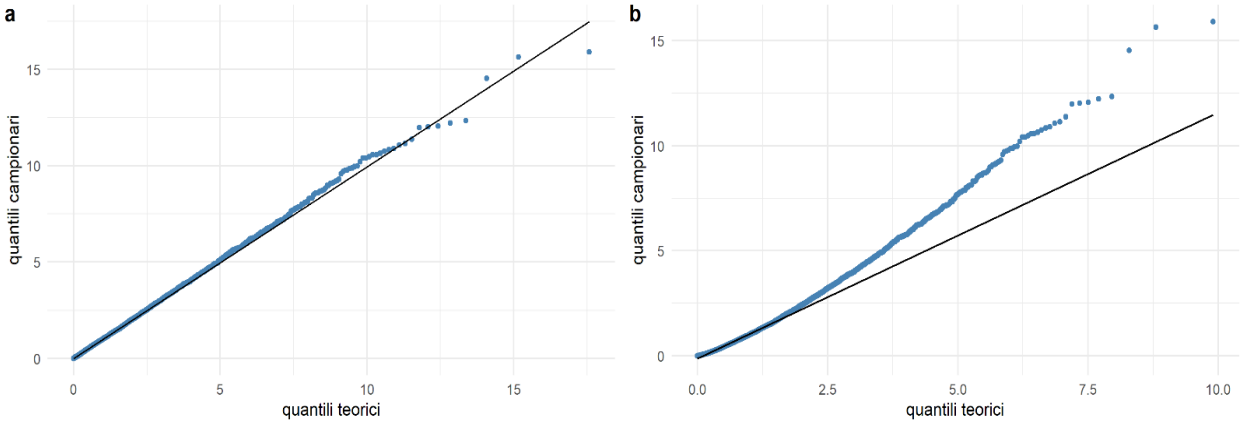

Figura 9: esempi di Q-Q plot. Nel pannello (a) sono raffigurati i quantili di un campione randomico generato da una distribuzione di Weibull vs i quantili teorici della stessa distribuzione; nel pannello (b) lo stesso campione è confrontato coi quantili teorici di una distribuzione esponenziale. Fonte: elaborazione dell'autore.

Il grafico rappresentato in Figura 9 è esemplificativo dell'utilità di un *Q-Q plot.* Come si può notare dal pannello (b), i quantili campionari (generati casualmente da una distribuzione di Weibull) vanno a divergere sempre di più da una distribuzione esponenziale man mano che ci si allontana dal "corpo" della distribuzione (angolo inferiore sinistro) e ci si avvicina alla coda (angolo superiore destro), e in particolare il grafico assume una curvatura verso l'alto rispetto ai quantili della distribuzione esponenziale: un grafico di questo tipo, dunque, ci suggerisce che i dati a disposizione provengono da una distribuzione *heavy-tailed*, quale è la Weibull quando il parametro di forma è compreso tra 0 e 1 (in questo esempio pari a 0.8).

L'ultimo tipo di grafico che vale la pena menzionare in questa sede è il *Mean Excess Plot.* Si tratta di un grafico spesso utilizzato in situazioni in cui si cerca di adattare un modello *heavy-tailed* ai dati, come spesso accade nell'analisi dei rischi operativi e *cyber*; esso si rivela molto utile, ad esempio, come "guida operativa" per la scelta della soglia nell'applicazione del metodo POT. Infatti, per un valore fissato della soglia *u*, la *mean excess function* di una distribuzione X è definita come la media condizionata dei valori di X che superano *u*, sulla condizione che X sia maggiore di *u*. In simboli,

$$e(u) = E[X - u | X > u]$$

La *sample mean excess function*, cioè la stima empirica di questa funzione, è data dunque da[62]

$$e_n(u) = \frac{\sum_{j=1}^{n} (x_j - u)_+}{\sum_{j=1}^{n} \mathbb{I}_{\{x_j > u\}}}$$

In breve, il *mean excess plot* è un grafico dei valori assunti da $e_n(u)$ al variare dei valori di *u*, ed osservandolo si possono trarre diverse conclusioni. Innanzitutto, a seconda del trend

[62] Notazione data da Chernobai, Rachev and Fabozzi (2012).

che assumono i punti del grafico è possibile determinare che tipo di distribuzione si sta osservando: un'inclinazione verso l'alto suggerisce che si è in presenza di una distribuzione *heavy tailed*, mentre un'inclinazione verso il basso indica l'esatto opposto; una linea orizzontale, invece, suggerisce che si è in presenza di una distribuzione esponenziale (i vari casi sono illustrati in figura 10). Un'altra informazione molto utile che è possibile ottenere riguarda il valore ottimale di *u*: se il grafico assume un'inclinazione positiva oltre un certo valore soglia, è possibile inferire che oltre quel valore i dati si distribuiscano secondo una GPD[63].

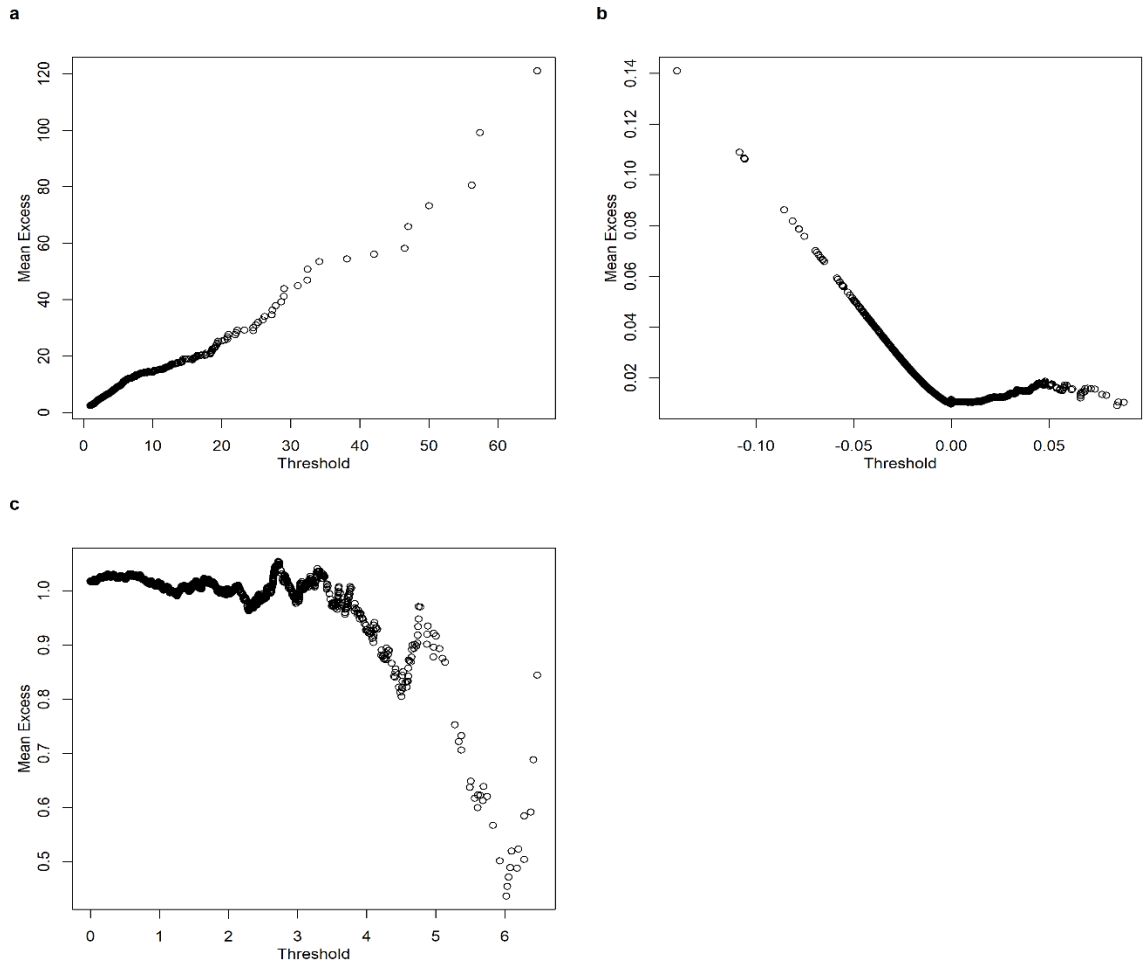

Figura 10: esempi di mean excess plot. I vari pannelli rappresentano i possibil casi descritti. In particolare (a) distribuzione heavy tailed; (b) distribuzione light tailed e (c) distribuzione esponenziale (andamento quasi orizzontale). Fonte: elaborazione dell'autore.

Come affermato in precedenza, le informazioni che è possibile desumere attraverso l'analisi di grafici come quelli appena descritti non sempre si rivelano corrette, o più semplicemente necessitano di validazione. Per accertare dunque la qualità del modello teorico ipotizzato è dunque necessario ricorrere a tecniche numeriche più formali quali i test d'ipotesi. Tipicamente, nell'ambito del *framework* delineato in questo capitolo, l'ipotesi nulla e alternativa che si vanno a testare sono le seguenti:

[63] Moscadelli (2004).

H_0: i dati seguono la distribuzione ipotizzata

vs

H_1: i dati non seguono la distribuzione ipotizzata

In questa sede descriveremo i tre test più comunemente utilizzati nella letteratura correlata: il test di Kolmogorov-Sminorv, il test di Anderson-Darling e quello di Cramér-von Mises.

Test di Kolmogorov-Smirnov. Esistono due varianti del test KS: la prima verifica che un dato campione provenga da una particolare distribuzione (*one-sample*) mentre la seconda verifica che due campioni diversi provengano dalla stessa distribuzione (*two-sample*). La statistica test utilizzata è la seguente:

$$KS = \sup_{x \in \mathbb{R}} \left| \widehat{F_n}(x) - F(x) \right| \qquad (2.20)$$

Dove $\widehat{F_n}(x)$ rappresenta la funzione di distribuzione empirica, mentre $F(x)$ rappresenta la funzione di distribuzione teorica: il test sostanzialmente va a calcolare la maggior distanza verticale tra le due distribuzioni. La (2.20) tuttavia non è una formula ideale a fini di calcolo, e in genere si tende ad utilizzare la seguente *computing formula*[64]:

$$KS = max\left\{ \sup_j \left(\frac{j}{n} - z_{(j)} \right), \sup_j \left(z_{(j)} - \frac{j-1}{n} \right) \right\} \qquad (2.21)$$

Intuitivamente, la statistica KS fornisce il massimo valore di discrepanza fra le due funzioni di ripartizione, ed è quindi agevole concludere che al crescere di tale valore sarà più probabile rifiutare l'ipotesi nulla; come regola operativa si guarda in genere al *p-value* fornito da tutti i *software* statistici: se tale valore è inferiore a 0.05, si rifiuta l'ipotesi nulla.

Il test di KS è sicuramente tra i più popolari, ma non è esente da difetti. Il principale svantaggio di questo tipo di test è dovuto al fatto che tende a sovrastimare il peso dei quantili più vicini alla mediana e di conseguenza a sottostimare il peso di quelli più lontani; in altre parole, il peso delle code verrà sottostimato dal test, fatto che rende il test poco ideale qualora si vada a testare il *fit* di modelli *heavy-tailed*[65].

Test di Anderson-Darling. Similarmente al test di Kolmogorov-Smirnov, il test AD va a misurare la discrepanza fra le funzioni di ripartizione empirica e teorica; a differenza del test

[64] Formula data da Chernobai, Rachev and Fabozzi (2012).
[65] Cruz, Peters and Shevchenko (2015).

KS, tuttavia, esso pone maggior peso sul *fit* delle code piuttosto che del corpo della distribuzione, ed è quindi maggiormente indicato quando si ritiene che i dati a disposizione siano *heavy-tailed*. La statistica utilizzata dal test è la seguente:

$$A^2 = n \int_t^u \frac{\left[\widehat{F}_n(x) - F(x)\right]^2}{F(x)[1 - F(x)]} f(x)dx \qquad (2.22)$$

Questa formula fornisce una media ponderata delle differenze quadratiche tra le funzioni di ripartizione[66]. La *computing formula* corrispondente è invece la seguente:

$$AD^2 = -nF(u) + n\sum_{j=0}^{k}\left[1 - \widehat{F}_n(y_j)\right]^2\{\ln[1 - F(y_{j+1})]\} + n\sum_{j=1}^{k}\widehat{F}_n(y_j)^2\left[\ln\frac{F(y_{j+1})}{F(y_j)}\right] \qquad (2.23)$$

Test di Cramér-Von Mises. Come gli altri due, anche il test di Cramér-Von Mises si basa sulla misura della distanza tra le funzioni di ripartizione. Si tratta di un test quadratico, che utilizza la seguente statistica test

$$W^2 = n\int_{-\infty}^{\infty}\left[\widehat{F}_n(x) - F(x)\right]^2 dF(x) \qquad (2.24)$$

Similmente al test A-D. Anche il test C-vM può essere effettuato come *one sample* o *two sample*, come già visto per il K-S.

[66] Panjer (2006).

Capitolo 3

Analisi Empirica

3.1 Introduzione all'analisi

In questo capitolo sono presentati i risultati dell'analisi condotta attraverso la metodologia del Loss Distribution Approach descritta nel paragrafo 2.4. Tale analisi è stata svolta partendo da uno dei dataset pubblici più diffusi in materia[67], il "Data Breach Chronology" (DBC) costruito e reso disponibile dalla Privacy Rights Clearinghouse (PRC), un'organizzazione no-profit statunitense che si occupa di protezione dei consumatori e della loro privacy. Si tratta di un dataset abbastanza ampio, contenente informazioni su oltre novemila *data breach* avvenuti tra il 2005 e il 2019 a discapito di organizzazioni americane; sono escluse tuttavia dal dataset tutte quelle fattispecie di diversa natura (descritte nel Capitolo 1) che vanno a costituire l'insieme dei *cyber risk*. Nonostante ciò, i *data breach* rappresentano sicuramente una delle maggiori componenti di tale insieme, e pertanto si ritiene che l'analisi effettuata sia sufficientemente rappresentativa del fenomeno.

In generale, è tuttavia azzardato affermare che i dati utilizzati in ricerche di questo tipo siano completi e "affidabili". Esistono infatti una serie di difficoltà nella raccolta di questi dati, anch'esse descritte nei capitoli precedenti, che allo stato attuale non sono state ancora risolte in maniera definitiva; tale problema è poi particolarmente evidente quando si va ad utilizzare dataset pubblicamente disponibili, che per loro natura registrano solamente eventi resi noti al pubblico attraverso i mezzi d'informazione. È dunque opportuno tenere a mente che i risultati conseguiti sulla base di queste premesse sono necessariamente illustrativi piuttosto che conclusivi; ciò vale anche per il lavoro di seguito presentato.

Nella prima parte di questo capitolo sarà dunque presentata un'analisi esplorativa del dataset DBC e delle sue caratteristiche, con particolare enfasi su di un sottoinsieme costituito dagli eventi riguardanti le organizzazioni appartenenti al settore finanziario; nella seconda parte invece sarà illustrata nel dettaglio la procedura adottata per la costruzione di un modello adeguato alla rappresentazione del *cyber risk,* prendendo come riferimento il sottoinsieme del dataset composto dalle imprese appartenenti al settore finanziario; il modello verrà poi applicato al dataset nella sua interezza. Infine, i risultati ottenuti dal modello saranno presentati, confrontati e commentati.

[67] Utilizzato, ad esempio, da Edwards, Hofmeyr and Forrest (2016).

3.2 Il dataset

Come accennato, il dataset *Data Breach Chronology* contiene informazioni su 9015 *data breach* subiti da organizzazioni statunitensi, pubbliche o private, appartenenti a diversi settori e in un arco di tempo che va dal 2005 al 2019. Le informazioni fornite dal DBC utili ai fini della nostra analisi sono essenzialmente 4:

- Anno dell'evento;
- Numero di *personal records* rubati nell'evento;
- Tipo di organizzazione (macrosettore di appartenenza);
- Tipo di attacco;

Nel dataset, ogni evento è categorizzato infatti anche in base al macrosettore di appartenenza e al tipo di attacco perpetrato, riassunti e descritti nella tabella 4.

Tipi di attacco		**Macrosettori**	
CARD: frodi perpetrate attraverso l'utilizzo di carte di debito/credito, ma che non coinvolgono *hacking* delle stesse (ad esempio *skimming*).	DISC: diffusione non intenzionale di informazioni sensibili, non dovuta ad *hacking* o perdita fisica.	BSF: settore finanziario e assicurativo.	MED: settore della sanità, dei fornitori di servizi medici e delle assicurazioni sanitarie.
HACK: *hacking* perpetrato da un terzo, ad esempio attraverso *malware*.	INSD: *insider action*, ad esempio dovuta a impiegati o fornitori.	BSR: settore retail (incluso online).	EDU: settore scolastico e universitario.
PHYS: perdita fisica di documenti.	PORT: perdita di dispositivi portatili come *laptop*, *hard drive*, *smartphone*, etc.	BSO: altri settori dell'economia	NGO: organizzazioni no-profit.
STAT: perdita, furto o accesso abusivo di dispositivi fissi, come computer o server.	UNKN: informazioni non sufficienti per la categorizzazione.	GOV: istituzioni governative o militari.	UNKN: informazioni non sufficienti per la categorizzazione.

Tabella 4: descrizione dei tipi di attacco e dei macrosettori utilizzati nel dataset DBC. Fonte: https://privacyrights.org/data-breaches.

In figura 11 sono riportate le frequenze assolute nel dataset dei tipi di attacchi, riportati con le etichette descritte in tabella 4, suddivisi per macrosettori, anch'essi riportati con le relative etichette. Da questo grafico è possibile ricavare alcune informazioni iniziali: il settore in assoluto più colpito è quello sanitario, seguito dal macrosettore generico 'BSO' e dal settore finanziario.

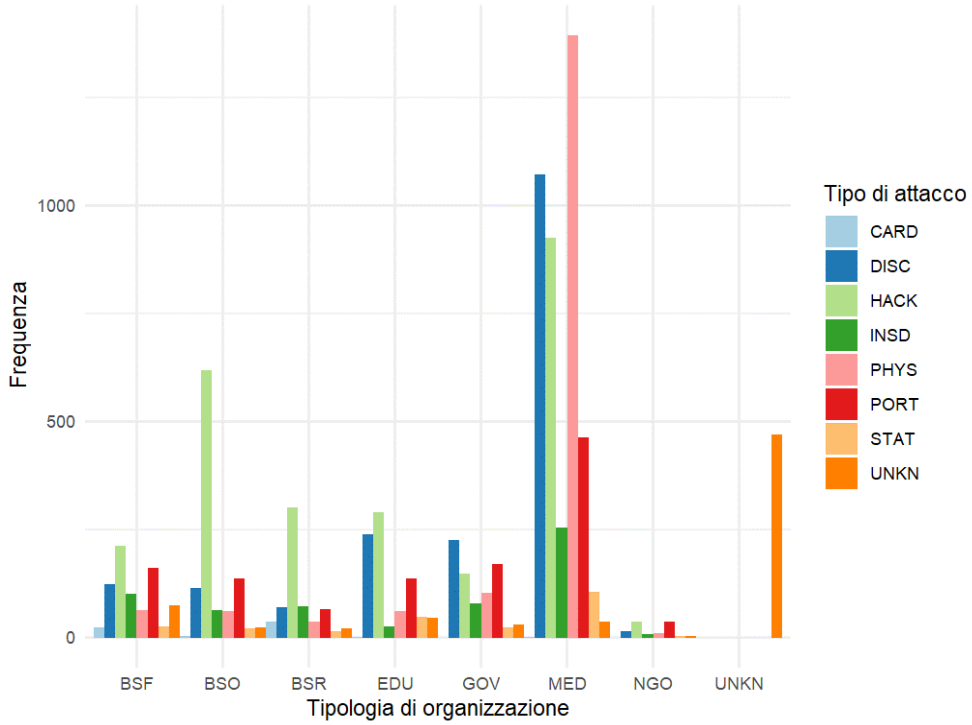

Figura 11: Tipi di cyber incident più diffusi per macrosettore nel dataset DBC.

Si tratta di dati che in effetti confermano la rilevanza anche a livello "sistemico" del *cyber risk*, come già si è discusso nel paragrafo 1.3, se si pensa alla centralità di cui il settore sanitario e quello finanziario godono nella vita sociale di ogni Stato moderno.

Un altro dato interessante da osservare riguarda poi la frequenza annuale degli attacchi, che sarà oggetto fondamentale dell'analisi svolta nei prossimi paragrafi. Come è possibile notare dal grafico riportato in figura 12, tra il 2016 e il 2017 si è assistito ad un significativo incremento di *data breach incidents*, in particolare attraverso l'*hacking* e la diffusione non intenzionale di dati sensibili; per quanto riguarda invece gli anni precedenti, è possibile vedere come, escluso l'anno iniziale (2005), non si ravvisa un aumento significativo di anno in anno, in linea con i risultati dei test più formali riportati in Edwards, Hofmeyr and Forrest (2016)[68].

[68] È importante sottolineare che nel dataset non sono inclusi dati relativi agli anni della pandemia da Covid-19 che, come già anticipato nel Capitolo 1, hanno visto un ulteriore sostanziale incremento nella frequenza dei *cyber incidents*, principalmente dovuto alla necessità del lavoro in remoto.

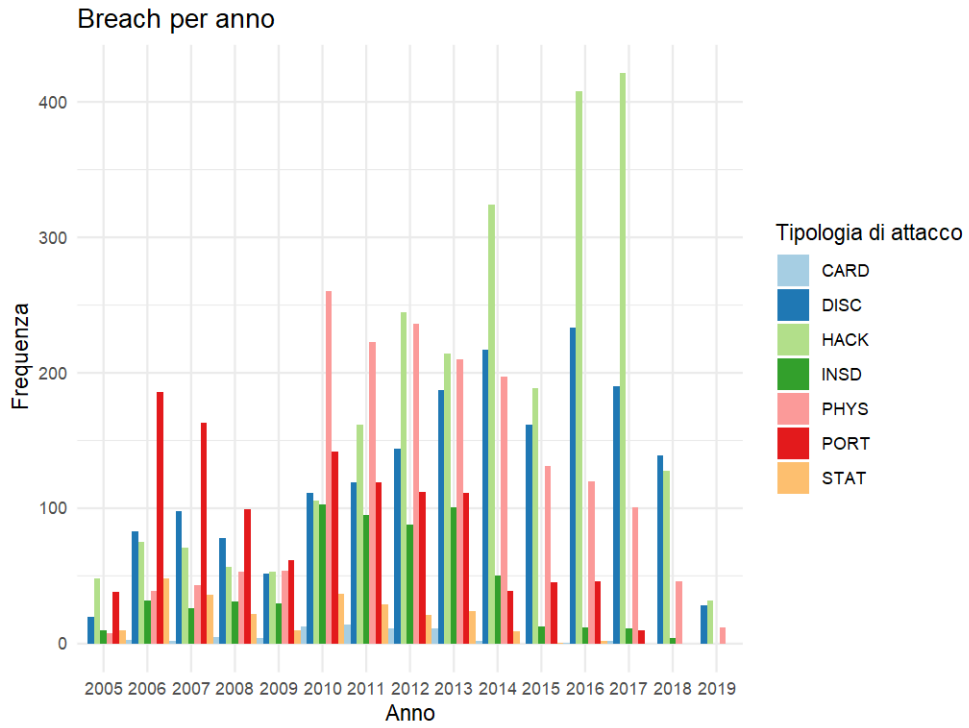

Figura 12: Frequenza annuale degli incidenti, suddivisi per tipologia, nel dataset DBC.

L'altra informazione fondamentale fornita dal dataset riguarda poi il numero di *records* persi o illegalmente sottratti da terzi per ogni attacco. Innanzitutto, è importante sottolineare come ben 2187 degli eventi registrati nel dataset (poco più del 24% delle osservazioni) non abbiano portato conseguenze, con nessun *record* perso; a tal proposito, si può pensare alla percentuale di eventi "nulli" (riportata in figura 13) come una *proxy* del grado di protezione di ogni macrosettore incluso nel dataset.

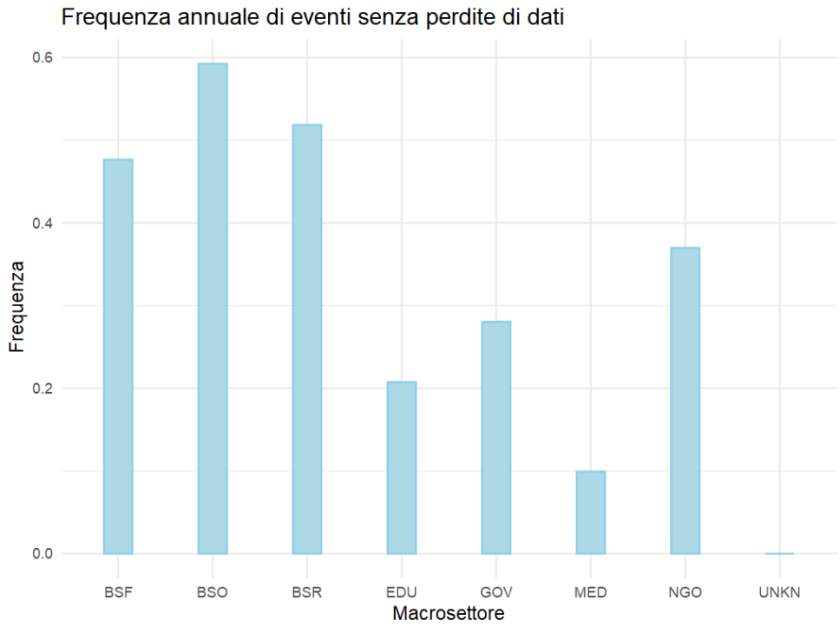

Figura 13: Percentuale di eventi "nulli" suddivisi per macrosettore.

Dalla figura notiamo che i settori meglio protetti risultano il settore generico-industriale (addirittura quasi 60% di eventi che non hanno comportato perdita di dati), quello retail (poco più del 50%) e il settore finanziario-assicurativo (poco meno del 50%). Quest'ultimo dato in particolare è rappresentativo della differenza nel grado di protezione, e di riflesso di investimenti in *cybersecurity*, tra il settore privato e quello pubblico; una differenza che diventa allarmante considerando che il settore sanitario, l'altro settore di rilevanza sistemica oltre a quello finanziario, risulta drammaticamente esposto in quanto riesce a "neutralizzare" solo il 10% degli incidenti su un totale di 4343 osservazioni.

La restante parte del dataset è costituita poi dagli eventi che hanno effettivamente portato ad una perdita di dati sensibili, le cui statistiche descrittive sono riportate in tabella 5 per il totale e per i 3 settori più colpiti.

Statistiche descrittive	Totale	Settore finanziario	Settore sanitario	Settore "business"
Numerosità campionaria	6822	411	3911	426
Minimo	1	1	1	2
Primo quartile	613	200	826	619
Mediana	2.000	2.000	2.000	5.750
Terzo quartile	10.000	27.000	7.752	81.620
Massimo	3.000.000.000	145.500.000	78.800.000	3.000.000.000
Media	1.522.632	1.566.457	63.691	18.307.912
Dev. Standard	41.960.690	11.312.621	1.309.715	165.319.323

Tabella 5: Statistiche descrittive del dataset DBC completo e dei tre settori più colpiti dai cyber incident, esclusi gli eventi con perdita nulla.

La prima particolarità che balza all'occhio è la forte asimmetricità e variabilità dei dati, come confermato anche dai valori di deviazione standard: sia per il dataset completo che per i tre sottoinsiemi, la differenza tra terzo quartile e valore massimo è ampissima. È questa un'informazione che conferma quanto affermato nel Capitolo 1 riguardo la "bi-modalità" di questo tipo di rischio: il corpo centrale dei dataset è formato dagli eventi più frequenti e a basso impatto, ma sono poi presenti anche eventi molto più rari e dall'impatto molto più devastante. Questa deduzione è confermata anche dall'istogramma del logaritmo dei dati riportato in figura 14; la forte variabilità e la presenza di *outlier* di diversi ordini di grandezza superiori alla media rende infatti impossibile la visualizzazione in scala naturale.

Un'altra informazione interessante è data dalla media di *record* persi per i 3 settori considerati: mentre il settore finanziario si mantiene sostanzialmente in linea con la media generale, il settore sanitario e quello generico si pongono, rispettivamente, sensibilmente al di sotto e al di sopra della media generale. Questo dato può essere interpretato come indicativo dell'attrattività che il settore business esercita sui *cybercriminali* rispetto a quello sanitario, che

di contro risulta molto più esposto ad eventi accidentali (categorie “PHYS” e “DISC”, vedi figura 11) che a minacce esterne, anche in virtù della differente natura dei dati sensibili detenuti.

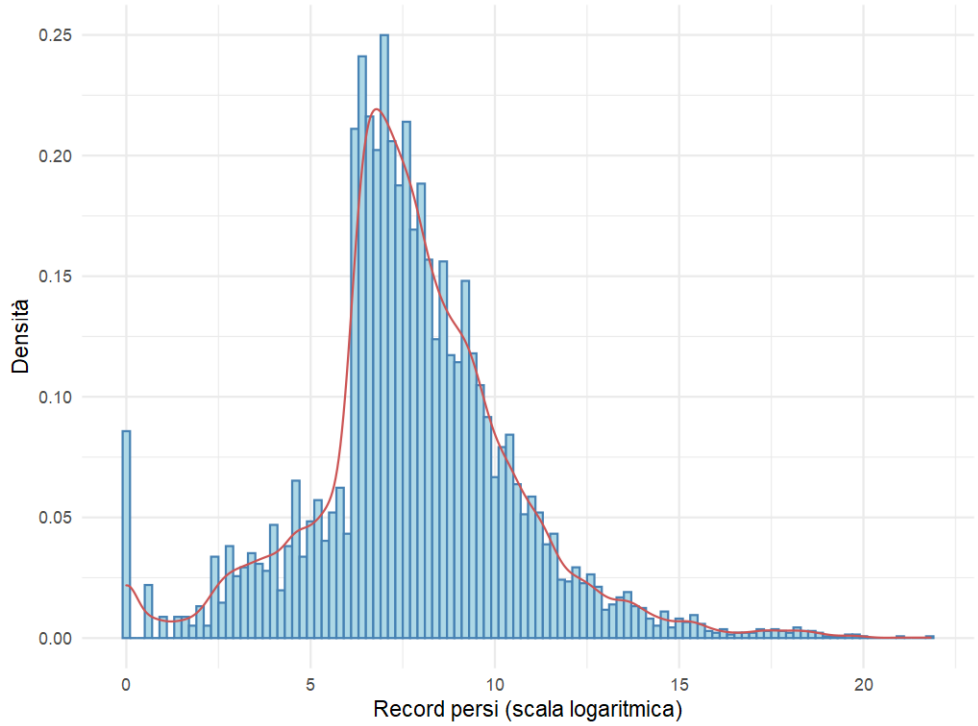

Figura 14: istogramma del logaritmo dei dati nel dataset DBC.

A conclusione di questa analisi esplorativa dei dati è necessario fare un’ultima osservazione che sarà poi approfondita in chiusura di capitolo. Come si sarà notato, il dataset non riporta alcuna informazione in merito alle perdite finanziarie associate ad ogni evento; è infatti un compito arduo stabilire una relazione diretta tra numero di dati persi o sottratti e perdita in termini monetari, in quanto al crescere del numero di *record* intervengono diversi fattori che influenzano, in positivo o in negativo, il costo sopportato dall’entità che subisce il *data breach*. L’analisi che segue, dunque, esprimerà un *Value-at-Risk* e un *Expected Shortfall* in termini di *record* persi o sottratti, e non in termini monetari; nel paragrafo 3.X sarà poi illustrato un metodo di conversione che è possibile impiegare per ottenere una misura monetaria delle perdite associate a questi eventi.

3.3 La costruzione del modello

L’analisi empirica di seguito presentata si fonda sull’applicazione del *Loss Distribution Approach* presentato nel Capitolo 2. Attraverso questo approccio è stato possibile dunque creare un modello teorico per la caratterizzazione del *cyber risk*; in realtà, poiché tale modello è stato calibrato sul dataset DBC, esso va a descrivere, più precisamente, un sottoinsieme di questa categoria di rischio, ossia il rischio collegato ai *data breach*. Inoltre, la calibrazione iniziale dei parametri e la scelta delle distribuzioni più adatte a descrivere i dati del modello è stata

effettuata sul sottoinsieme del dataset comprendente i soli eventi che hanno colpito organizzazioni classificate nel settore finanziario. Nonostante ciò, si tratta di un modello che può essere facilmente esteso e ricalibrato.

Il modello può essere riassunto e descritto dall'equazione (2.11), di seguito riportata:

$$Z_j = \sum_{i=1}^{N_j} X_i^{(j)} = X_1^{(j)} + X_2^{(j)} + \cdots + X_{N_j}^{(j)}$$

dove si ricorderà che Z_j indica la distribuzione aggregata delle perdite per la j-esima combinazione tra tipo di evento/linea di business, $X_i^{(j)}$ indica la i-esima perdita associata ad un data breach per la j-esima combinazione, e N_j indica la distribuzione di frequenza degli eventi. Nel prosieguo del testo, faremo l'ipotesi semplificatrice che il dataset a nostra disposizione contenga dati relativi ad una singola combinazione evento/business line, e dunque elimineremo l'indice j:

$$Z = \sum_{i=1}^{N} X_i = X_1 + X_2 + \cdots + X_N \qquad (3.1)$$

I parametri delle distribuzioni di severity (X_i) e di frequenza (N) saranno stimati attraverso il metodo della massima verosimiglianza (*maximum-likelihood estimation*) e la relativa bontà di adattamento ai dati sarà valutata attraverso le varie tecniche illustrate nel Capitolo 2; le ipotesi che saranno testate sono riassunte in tabella 6. Dopo aver selezionato la combinazione più soddisfacente, la distribuzione aggregata Z sarà calcolata attraverso un algoritmo di simulazione numerica (metodo Monte Carlo), descritto in dettaglio nel prosieguo del capitolo. Ottenuta la distribuzione aggregata sarà possibile infine calcolare le misure di rischio (*VaR* ed *Expected Shortfall*) come output finale del modello (vedi figura 15).

Frequenza	*Severity*
1. La frequenza degli eventi si distribuisce come una Poisson (λ); 2. La frequenza degli eventi si distribuisce come una somma di Poisson (λ_1+λ_2). Le distribuzioni marginali si riferiscono al "corpo" e alla "coda"; 3. La frequenza degli eventi si distribuisce come una Binomiale Negativa (*r, β*); 4. La frequenza degli eventi si distribuisce secondo una mistura di due distribuzioni BN, relative rispettivamente al "corpo" e alla coda dei dati.	1. La severità delle perdite si distribuisce secondo una distribuzione lognormale; 2. La severità delle perdite si distribuisce come una distribuzione di Weibull; 3. La severità delle perdite segue una distribuzione Esponenziale; 4. La severità delle perdite è descritta da un modello misto, dove il "corpo" della distribuzione è modellato secondo una lognormale o una Weibull, e la coda secondo una *Generalized Pareto* (GPD).

Tabella 6: ipotesi sulle distribuzioni di frequenza e di severity dei data breach.

In alcuni casi, il dataset è stato scomposto sulla base di un valore soglia *u*, dividendo gli eventi tra "corpo", cioè con perdita di *records* inferiore alla soglia fissata, e "coda", relativa agli eventi con perdita superiore a *u*. Le distribuzioni di frequenza e *severity* sono state stimate separatamente per entrambi i sotto-dataset per testare alcune di queste ipotesi; questo approccio prende il nome di *Piecewise-Defined Loss Distribution Approach* e rappresenta una variante del classico LDA[69].

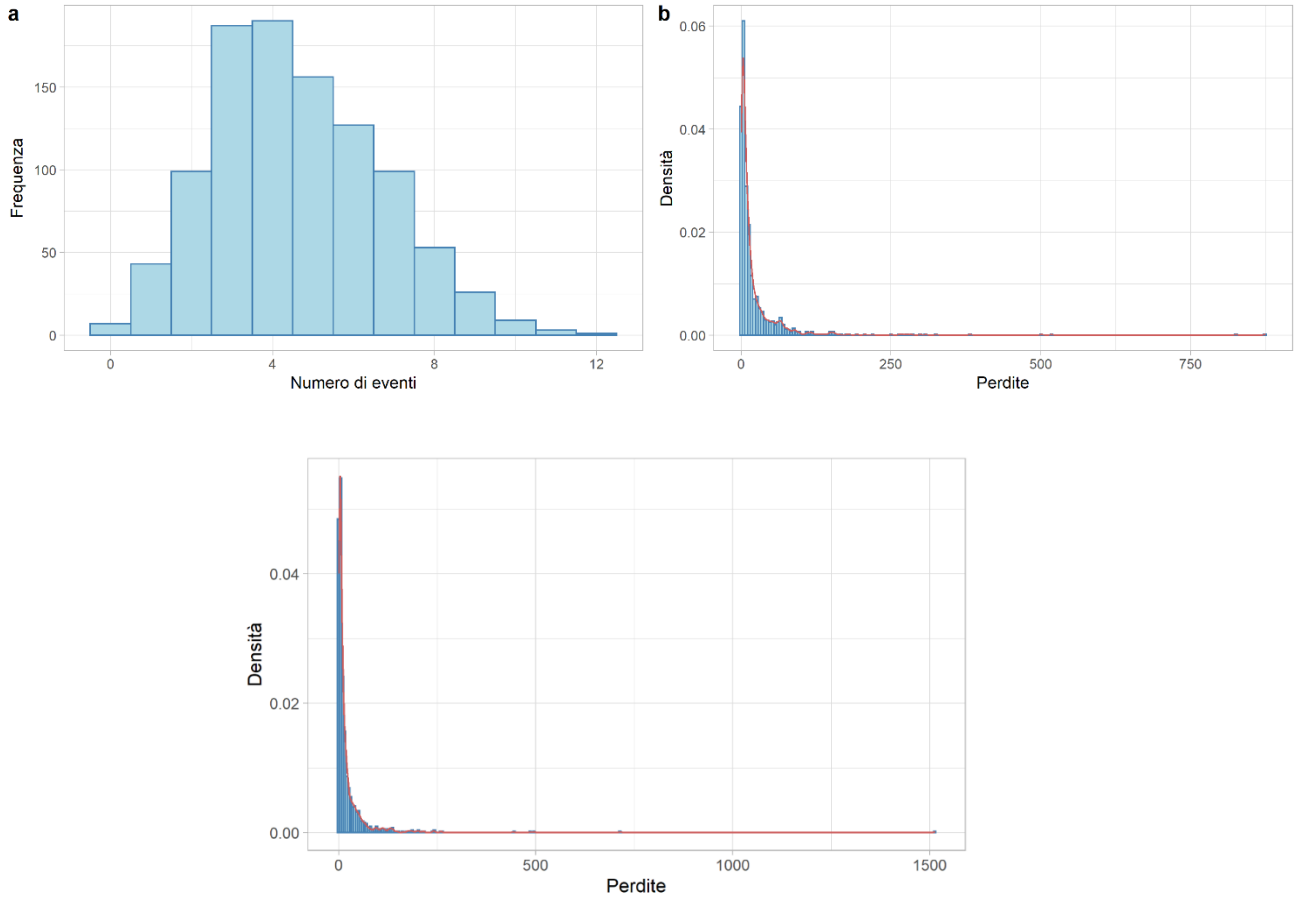

Figura 15: Esempio grafico del Loss Distribution Approach. Nei pannelli (a) e (b) sono raffigurate rispettivamente una distribuzione "tipo" di frequenza e di impatto. Nel pannello in basso invece è raffigurata la distribuzione aggregata delle perdite ottenuta combinando le due distribuzioni. Fonte: elaborazione dell'autore.

In figura 16 è presente il *Mean Excess Plot* dei dati, che segnala chiaramente che si ha a che fare con una distribuzione *heavy-tailed*; per una miglior visualizzazione, nel grafico raffigurato nel pannello (a) della figura sono state omesse le prime venti osservazioni di maggior entità nel dataset, mentre nel pannello (b) sono state omesse le prime trenta osservazioni ed è quindi una versione "zoomata" del grafico.

[69] Un'applicazione pratica di questo approccio è presentata in Li, Feng and Chen (2009).

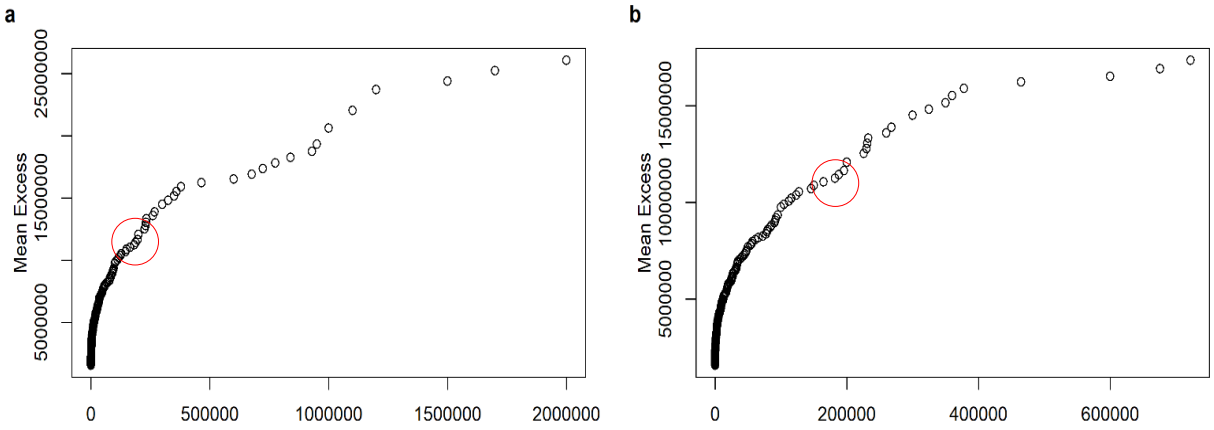

Figura 16: mean excess plot del dataset "finanziario". Il grafico presenta un'inclinazione positiva segnalando un comportamento heavy-tailed dei dati. In rosso è cerchiato il punto in cui il grafico assume un'inclinazione linearmente positiva.

Il punto del grafico cerchiato in rosso individua il valore candidato ad essere selezionato come soglia *u* (corrispondente a circa 200.000 *records* persi); infatti, oltre questo punto il grafico assume un andamento lineare e positivo, segnale di un comportamento paretiano nella coda della distribuzione.

3.3.1 La scelta della distribuzione di frequenza

Una volta individuato il valore soglia *u*, pari a 200.000, il passo successivo consiste nella stima della distribuzione di frequenza. L'obiettivo che si persegue attraverso la stima di tale distribuzione è quello di ottenere un modello teorico che dia la probabilità che un determinato numero di eventi si verifichi durante un dato periodo di tempo, nel nostro caso un anno. Attraverso l'analisi dei dati passati si cerca dunque di "prevedere" quale potrà essere l'evoluzione del fenomeno dal punto di vista "numerico", cioè della frequenza degli eventi futuri.

Come si ricorderà dal Capitolo 2, le due distribuzioni tipicamente più utilizzate per la modellizzazione della frequenza nell'ambito del rischio operativo, e per estensione del *cyber risk*, sono la distribuzione di Poisson (λ) e la distribuzione Binomiale Negativa (r, β). Inizialmente si è dunque proceduto alla stima dei parametri di tali distribuzioni sull'intero dataset e alla valutazione della bontà di adattamento ai dati; in figura 17 e 18 sono riportati, rispettivamente, il confronto tra le funzioni di probabilità (a) e di ripartizione (b) empiriche e teoriche insieme ai relativi *Q-Q Plot* (c).

Dall'osservazione dei due set di grafici è possibile iniziare ad apprezzare il fatto che la distribuzione BN si adatti meglio ai dati a nostra disposizione. Si tratta in realtà di un risultato che non dovrebbe sorprendere; la BN è infatti una distribuzione la cui varianza è sempre

maggiore della media e che si presta dunque meglio per modellare eventi caratterizzati da un'alta volatilità, che la Poisson non riuscirebbe a catturare adeguatamente[70].

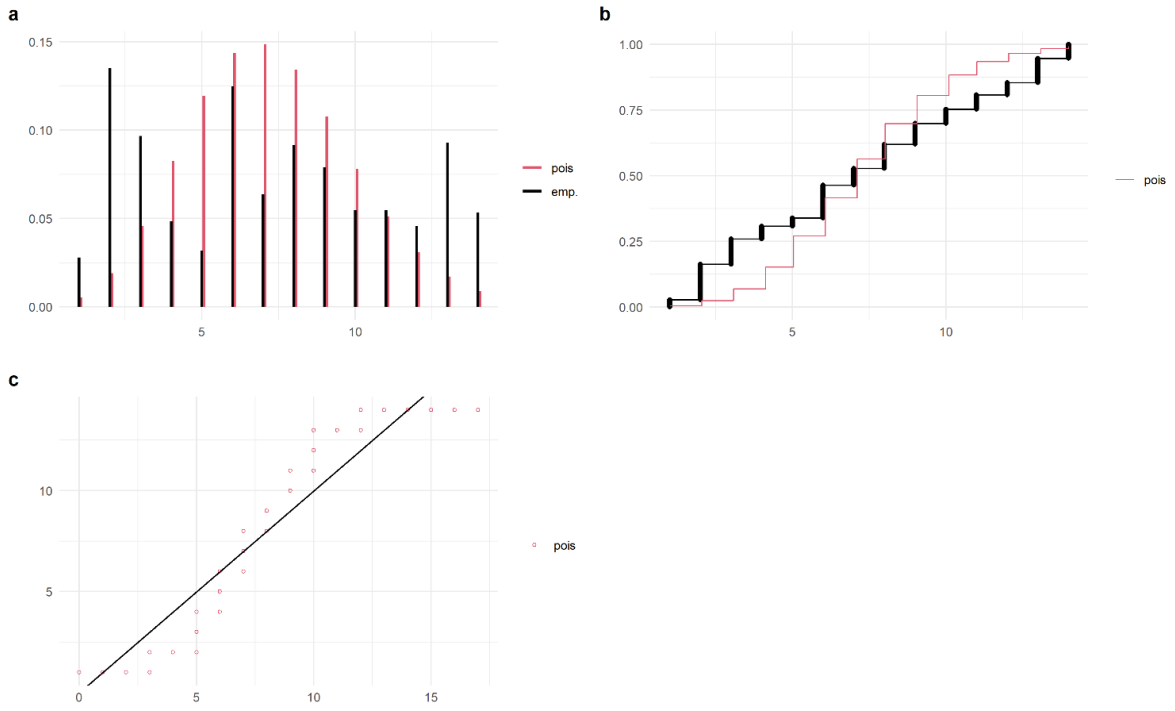

Figura 17: grafici di confronto tra la distribuzione di frequenza empirica e Poisson stimata.

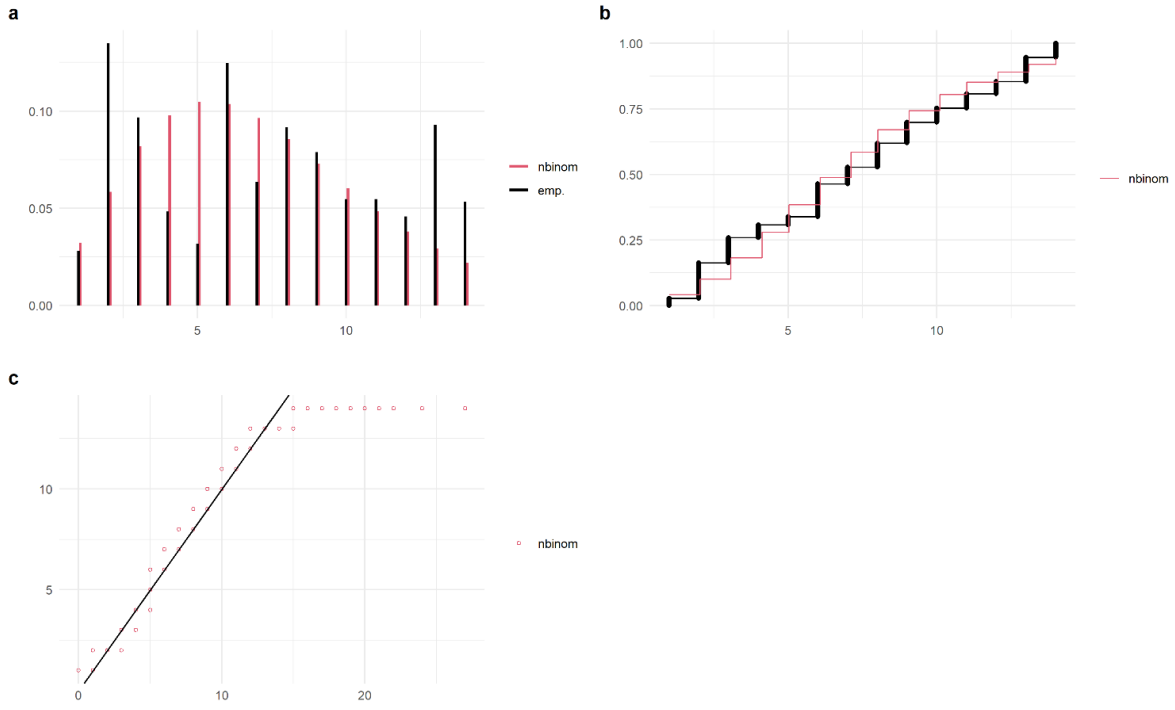

Figura 18: grafici di confronto tra la distribuzione di frequenza empirica e Binomiale Negativa stimata.

[70] Parodi (2014).

Una rapida analisi dei grafici mostra che il grafico della funzione di ripartizione (CDF) della BN segue più da vicino quello della distribuzione empirica dei dati, e anche il relativo *Q-Q plot* conferma la presenza di un *fit* migliore, che tuttavia non significa perfetto; il grafico di comparazione tra le funzioni di probabilità (pannello a) mostra infatti come anche la BN tenda a sovrastimare o sottostimare la frequenza degli eventi in alcuni casi. Le impressioni sono confermate dai dati riportati in tabella 7, che contiene la stima dei parametri delle distribuzioni (con i relativi *standard error*) e i *p-value* relativi ai test sulla bontà di adattamento.

Distribuzione teorica	**Parametri stimati**		*Goodness-of-Fit (p-value)* **A - D**	**C - VM**
Poisson	λ = 7,23028 (0,09591053)		0,0006019	0,029
Binomiale Negativa	μ = 7,229802 (0,1479136)	size = 5,243556 (0,4813624)	0,2121	0,1788

Tabella 7: Stima dei parametri e relativi test di bontà di adattamento per le distribuzioni di frequenza.

Sia il test di Anderson-Darling che quello di Cramér-von Mises confermano che il *fit* della Binomiale Negativa è sostanzialmente migliore; entrambi i *p-value* relativi alla distribuzione di Poisson, infatti, portano a rifiutare l'ipotesi nulla di appartenenza.

Seguendo poi l'intuizione di Li, Feng and Chen (2009), si è proceduto alla stima delle distribuzioni di frequenza per i due sottoinsiemi del campione, l'uno relativo alle perdite più frequenti e di minor entità (inferiori alla soglia *u*) e l'altro relativo alle perdite più rare ma di maggior impatto (superiori alla soglia fissata). Nel lavoro citato gli autori ricorrono per semplicità ad una delle proprietà della distribuzione di Poisson, ossia che

$$Pois(\lambda_1) + Pois(\lambda_2) = Pois(\lambda_1 + \lambda_2)$$

e cioè, la somma di due distribuzioni di Poisson è ancora una Poisson con parametro pari alla somma dei lambda delle distribuzioni marginali. Tuttavia, come appena visto, la distribuzione di Poisson non si adatta bene ai dati a disposizione. Si è dunque provato a modellare la frequenza degli eventi attraverso una mistura di due distribuzioni Binomiali Negative, relative rispettivamente alle due tipologie di eventi. Una mistura di distribuzioni (*mixture distribution*) è una variabile casuale a sua volta, la cui funzione di probabilità è data dalla media ponderata delle funzioni di probabilità delle variabili che la compongono; in questo caso, come pesi sono state utilizzate le grandezze relative dei due sotto-campioni sul totale. Si ha infatti motivo di ricorrere all'utilizzo di una *mixture* quando si ritiene che una popolazione di valori sia composta

da più sottopopolazioni ciascuna definita da una propria distribuzione. I risultati dei test statistici per questa procedura sono riassunti in tabella 8.

Modello	Parametri stimati		A - D	C - VM
Binomiale Negativa (corpo)	μ = 7,346645 (0,1522009)	size = 5,615495 (0,5470564)	0,06369	0,04825
Binomiale Negativa (coda)	μ = 5,648391 (0,5466911)	size = 3,041547 (0,9208367)	0,4967	0,5203
Mistura			0,2112	0,1893

Tabella 8: stima dei parametri per i due sotto-campioni e della mistura di Binomiali Negative e test di Goodness-of-fit.

Come si vedrà più avanti, gli output finali del modello non sono significativamente diversi a seconda che si scelga di utilizzare la distribuzione "semplice" o la mistura di BN ed entrambe rappresentano una scelta ragionevole dal punto di vista statistico[71].

3.3.2 La scelta della distribuzione di *severity*

Anche per quanto riguarda la selezione di un modello adeguato a descrivere l'entità delle perdite, la scelta iniziale da fare è stata tra l'adottare una semplice distribuzione per modellare l'intero dataset a nostra disposizione o cercare di costruire un modello più complesso, composto da due distribuzioni, similmente a quanto già descritto in merito alla distribuzione di frequenza. Data la natura estremamente *over-dispersed* ed asimmetrica dei dati nel campione, e il fatto che nessuno dei modelli più elementari (log-normale, Weibull, e così via) sarebbe riuscito a catturare adeguatamente il comportamento *heavy-tailed* degli stessi (evidenziato già attraverso il *Mean Excess Plot* in figura 16), la scelta di un modello più complesso è stata quasi immediata. La nostra distribuzione di *severity*, dunque, è una distribuzione congiunta (*spliced*) nel punto *u* la cui funzione di ripartizione assume la forma:

$$F(x) = \begin{cases} F_1(x), & 0 \leq x \leq u \\ F_1(u) + \big(1 - F_1(u)\big) G_{u,\xi,\beta}(x), & x > u \end{cases} \qquad (3.2)$$

Il modello dell'entità delle perdite è dunque composto da due distribuzioni congiunte nel punto u, $F_1(x)$ e $G_{u,\xi,\beta}(x)$ che descrivono rispettivamente il corpo e la coda della distribuzione "totale"; in virtù di ciò, è comunque opportuno effettuare una valutazione sulle distribuzioni più semplici per poter selezionare quella più adatta a modellare gli eventi di

[71] È opportuno sottolineare che i *p-value* dei test non sono misure sulla base delle quali è possibile ordinare i migliori *fit*. Essi consentono esclusivamente di *non rifiutare* (e non di accettare in via definitiva) l'ipotesi nulla che i dati provengano da una data distribuzione ad un dato livello di significatività.

impatto minore. Le ipotesi testate sono già state presentate in tabella 6; la prima ad essere valutata è stata la distribuzione esponenziale per una ulteriore e definitiva conferma della natura *heavy-tailed* dei dati, come si può vedere dal confronto tra la funzione di ripartizione empirica e teorica e dal *Q-Q plot* in figura 19.

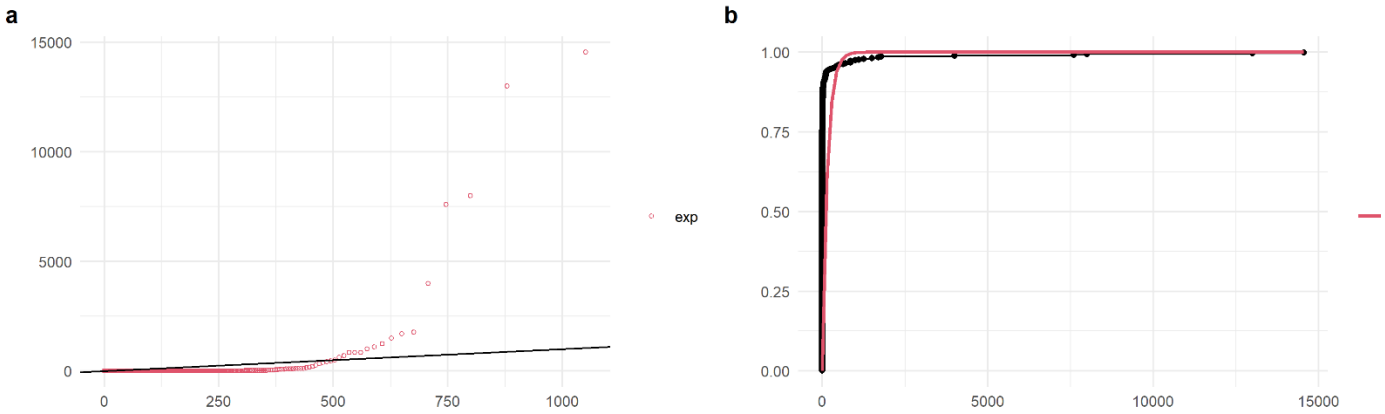

Figura 19: (a) Q-Q plot e (b) confronto tra la ECDF e la funzione di ripartizione della distribuzione esponenziale stimata. I dati sono stati ridimensionati di un fattore pari a 10.000 per permettere al software di stimare il parametro Θ. Dalla curvatura verso l'alto del Q-Q plot si nota chiaramente la natura heavy-tailed dei dati, confermata dal fatto che la coda della funzione di ripartizione empirica è sub-esponenziale.

Le distribuzioni considerate per modellare la parte al di sotto di *u* del dataset sono dunque la log-normale e la Weibull; in tabella 9 sono riportati i risultati della stima e dei test condotti.

Distribuzione teorica	Parametri stimati		*Goodness-of-Fit (p-value)*		
			A-D	K-S	C-VM
Log-Normale	μ = 7,919441 (0,1805038)	σ = 3,659378 (0, 1276354)	0,6708	0,1516	0,5878
Weibull	α = 0,251704 (0,008324)	β = 17.347,78 (2.150,86)	0,2767	0,0000 366	0,2188

Tabella 9: stima dei parametri e test di bontà di adattamento per il "corpo" della distribuzione di severity.

Nessuno dei test ci porta a rifiutare l'ipotesi nulla di coerenza con una delle due distribuzioni ipotizzate, fatta eccezione per il test K-S sulla distribuzione di Weibull; bisogna dunque fare affidamento su tecniche differenti per poter giungere ad una scelta adeguata. In figura 20 sono illustrati i *Q-Q plot* di entrambi i modelli, mentre in figura 21 sono confrontati i grafici delle funzioni di ripartizione empiriche e teoriche[72].

[72] Nell'osservazione di entrambe le tipologie di grafici è opportuno concentrarsi solo sul "corpo" dei dati in quanto le code saranno parametrizzate diversamente attraverso il metodo *Peaks-over-Threshold.*

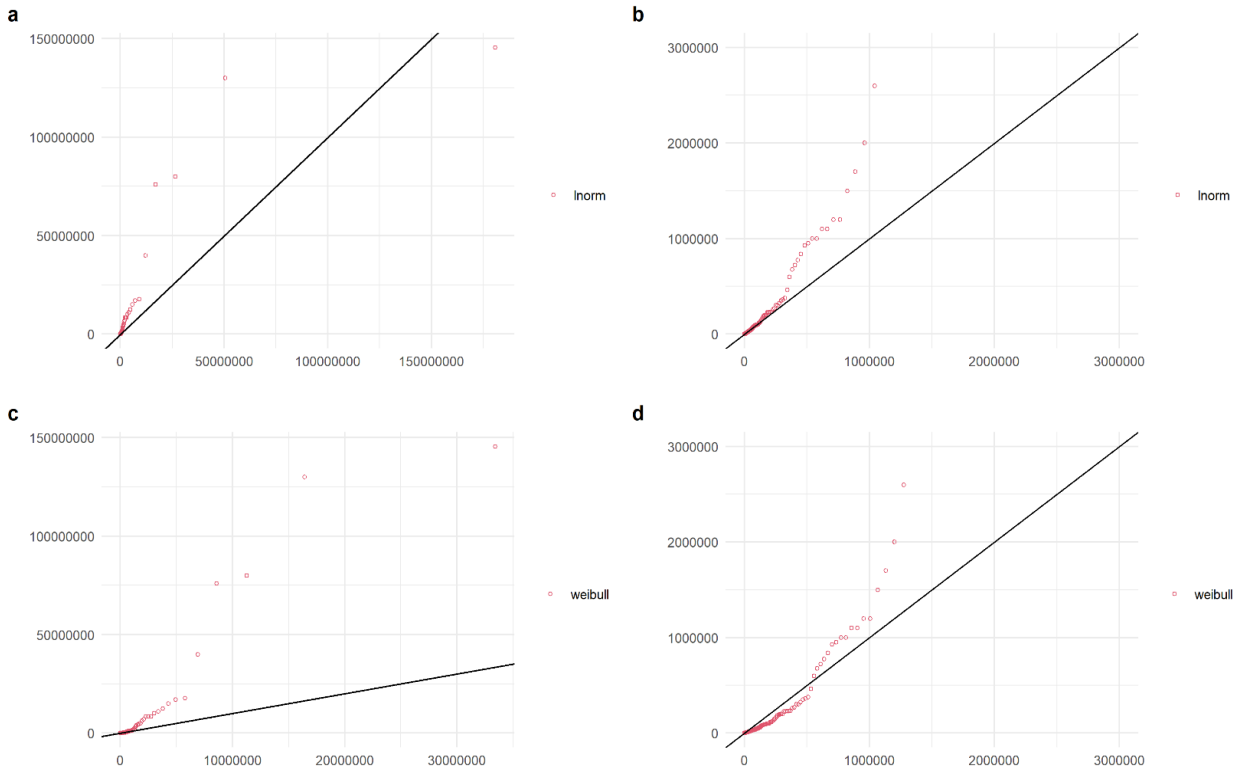

Figura 20: Q-Q plot per il fit lognormale (a) e Weibull (c). Nei pannelli (b) e (d) sono raffigurate versioni "zoomate" dei rispettivi grafici (troncati a x = 3.000.000). Quantili teorici sull'asse delle ascisse.

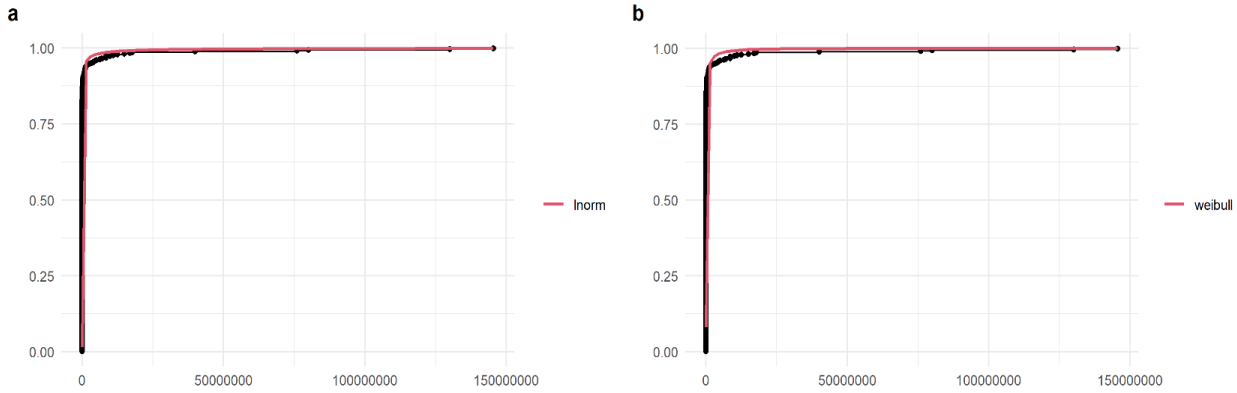

Figura 21: Grafico della funzione di ripartizione per il fit lognormale (a) e Weibull (c). Nei pannelli (b) e (d) sono raffigurate versioni "zoomate" dei rispettivi grafici (troncati a x = 3.000.000).

Dai grafici, la distribuzione log-normale sembra fornire un *fit* leggermente migliore nei dati che vanno a costituire il corpo della distribuzione empirica, impressione confermata anche dai criteri di informazione Bayesiano (BIC) e di Akaike (AIC) [73], riassunti per entrambi i modelli in tabella 10.

[73] L'AIC e il BIC sono criteri di selezione dei modelli parametrici stimati attraverso il metodo della massima verosimiglianza (MLE). Dato un set di modelli, essi forniscono una stima della qualità di ognuno di essi; in generale, il modello con i valori AIC e BIC più bassi tra quelli stimati risulta quello di miglior qualità.

Modello	AIC	BIC
Log-Normale	8.746,52	8.754,56
Weibull	8.862,97	8.871,01

Tabella 10: criteri d'informazione e di selezione dei modelli stimati.

Con le informazioni ottenute fino a questo punto, dunque, la scelta per il modello del "corpo" della distribuzione di *severity* ricade sulla distribuzione log-normale; è un risultato che non sorprende, poiché già dall'istogramma del logaritmo dei dati (figura 14) è possibile notare una densità approssimativamente Gaussiana[74]. Come risulta evidente anche dai grafici, tuttavia, il *fit* nella coda dei dati risulta insufficiente.

Per stimare la "restante parte" del modello di *severity* si è fatto ricorso al metodo POT della *Extreme Value Theory* illustrato nel Capitolo 2 e si è così pervenuti ad una *spliced distribution* i cui parametri sono riassunti in tabella11. Non possiamo rifiutare l'ipotesi di appartenenza dei dati alla distribuzione mista stimata, ma guardando ai grafici in figura non si rileva un apprezzabile miglioramento del *fit.*

Parametri Stimati		*Goodness-of-Fit (p-value)*		
Log-Normale	*Generalized Pareto*	A-D	K-S	C-vM
$\mu = 7{,}8308193$ $\sigma = 3{,}4816809$	$\beta = 23.323.229$ $\xi = 0{,}3000122$	0,774	0,308	0,704

Tabella 11: parametri stimati per il modello spliced Log-Normale + GPD con relativi test di Goodness-of-Fit.

[74] Si ricorda infatti dal Capitolo 2 che, se un campione di dati si distribuisce secondo una log-normale, allora il logaritmo degli stessi dati si distribuirà secondo una distribuzione Normale (o Gaussiana).

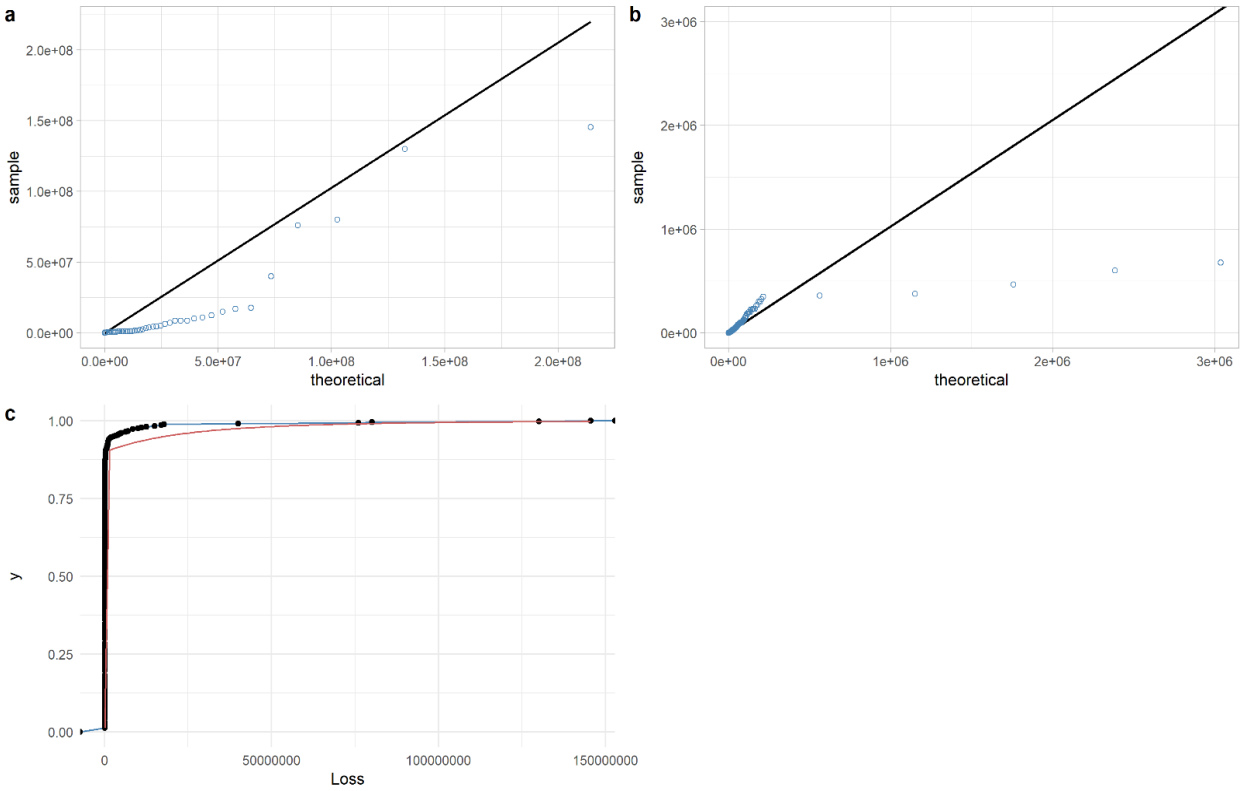

Figura 22: (a) Q-Q plot del modello misto stimato e (b) sua versione zoomata (assi limitati a 3.000.000). Nel pannello (c) il confronto tra le funzioni di ripartizione.

Il modello stimato, a livello grafico, non sembra adattarsi adeguatamente ai dati; dal grafico delle funzioni di ripartizione in particolare si nota una "spaccatura" molto marcata, presumibilmente localizzata nel punto in cui è stata fissata la soglia *u*.

Decidiamo dunque di imporre alla funzione di stima un vincolo di continuità nel punto *u* riducendo il parametro β a:

$$\beta_u = \frac{1 - H(u|\boldsymbol{\theta})}{h(u|\boldsymbol{\theta})}$$

dove con $H(u|\theta)$ indichiamo la funzione di ripartizione del "corpo" della distribuzione (e con h minuscola la relativa funzione di densità), mentre con $\boldsymbol{\theta}$ è indicato il vettore dei parametri della stessa[75]. La stima dei parametri e della bontà di adattamento del modello risultante dall'applicazione del vincolo è riportata in tabella 12; in figura 23 sono invece presenti i relativi grafici diagnostici.

[75] Hu and Scarrot (2018).

Parametri Stimati (modello con *continuity constraint*)		*Goodness-of-Fit (p-value)*		
Log-Normale	*Generalized Pareto*	A-D	K-S	C-vM
$\mu = 7{,}850566$ $\sigma = 3{,}516027$	$\beta = 408.955{,}2$ $\xi = 2{,}267518$	0,752	0,313	0,679

Tabella 12: parametri stimati per il modello spliced Log-Normale + GPD con vincolo di continuità e relativi test di Goodness-of-Fit.

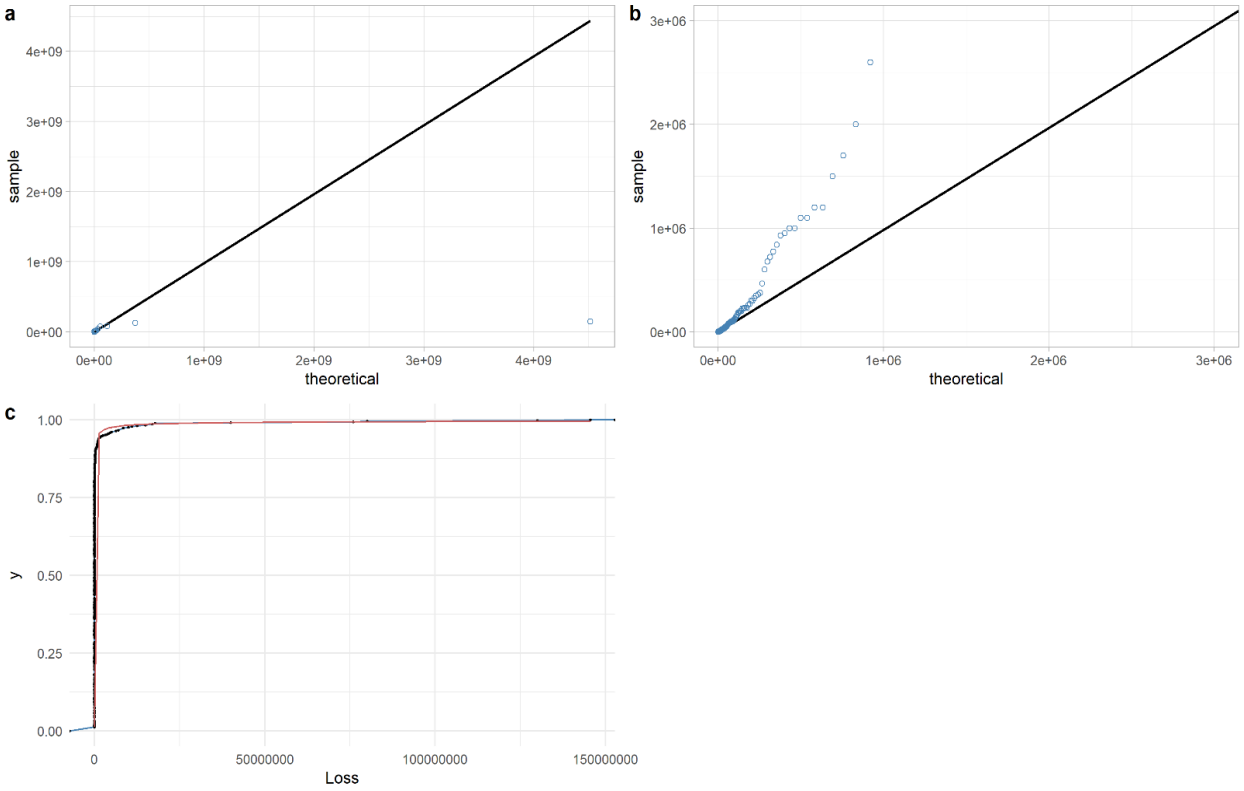

Figura 23: (a) Q-Q plot con relativa versione zoomata (b) del modello con vincolo di continuità. Nel pannello (c) il confronto tra le funzioni di ripartizione.

Dai grafici, e in particolare dal confronto tra funzioni di ripartizione, il *fit* del modello modificato risulta sensibilmente migliore, in particolare nell'area del punto u. Nonostante ciò, come si vedrà più avanti, i valori risultanti dall'adozione di questo modello saranno sensibilmente più alti (e probabilmente sovrastimati) rispetto a quelli risultanti dal modello senza vincolo.

3.3.3 La distribuzione aggregata delle perdite

Il passo conclusivo per la costruzione del modello consiste nell'aggregazione delle distribuzioni di frequenza e di *severity* stimate fino a questo punto, in modo da ottenere la distribuzione aggregata di probabilità delle perdite su un orizzonte temporale di un anno, Z. Si tratta di un'operazione che, in generale, può essere svolta in diversi modi più o meno efficienti; infatti, quando non è possibile ricorrere a formule analitiche ben definite, come in questo caso, si può fare affidamento su algoritmi numerici per ottenere un'approssimazione della

distribuzione aggregata. Tra questi, il metodo più diffuso è sicuramente il metodo Monte Carlo[76], che consiste nella simulazione di *n* scenari sulla base delle distribuzioni di frequenza e di *severity*, da "aggregare" successivamente in modo da ottenere una stima della nostra distribuzione d'interesse. La popolarità di questo metodo è da imputare principalmente alla sua flessibilità e semplicità concettuale: aumentando il numero di simulazioni, tendenzialmente, migliorerà anche la precisione della stima poiché si andranno a considerare un numero maggiore di scenari possibili, cosicché i parametri stimati convergano ai reali parametri della distribuzione all'aumentare di *n*[77].

A fronte di questa semplicità di utilizzo e di implementazione, il metodo Monte Carlo presenta alcuni difetti non banali: in primis, il numero di scenari da generare per ottenere un sufficiente grado di precisione non può essere stabilito a priori e dev'essere ottenuto attraverso un processo di *trial-and-error*; inoltre, a causa del fatto che il grado di precisione della stima è proporzionale ad *n*, il metodo risulta anche particolarmente "lento" e oneroso in termini computazionali[78].

L'algoritmo utilizzato in questo lavoro è riassumibile nei seguenti passi:

1. Data la stima dei parametri della distribuzione di frequenza, da questa è estratto un numero casuale x rappresentativo del numero di eventi nel *j*-esimo anno;
2. Dato il numero di eventi x, per ognuno di questi viene simulata una perdita dalla relativa distribuzione di *severity* stimata in precedenza. Ad esempio, supponendo che x = 5, verranno simulate cinque perdite di diversa entità;
3. Le perdite calcolate allo step 2 vengono sommate per ottenere la perdita cumulata annuale relativa all'anno *j*.
4. Si ripete il procedimento partendo dal punto 1 per un numero *n* di volte, rappresentative dei vari scenari.

Ripetendo la procedura un numero sufficiente di volte, si otterrà una stima numerica della distribuzione aggregata delle perdite annuali descritta dall'equazione (3.1) e da questa sarà poi possibile stimare le relative misure di rischio. Quanto detto risulterà più chiaro osservando l'esempio in tabella 13.

[76] Altri metodi molto diffusi includono, ad esempio, l'algoritmo di Panjer o il metodo FFT (Fast Fourier Transformation).
[77] È noto che nella stima dei requisiti di capitale via metodo Monte Carlo, l'accuratezza della stima cresce proporzionalmente alla radice quadrata del numero di scenari ($\sqrt{n}$). Fonte: Greselin, Piacenza and Zitikis (2019).
[78] Parodi (2014).

		Ammontare singole perdite							
Scenario	**N° perdite annue**	**#1**	**#2**	**#3**	**#4**	**#5**	**#6**	**#7**	**Totale**
#1	4	1.000	13.000	457.000	750				471.750
#2	5	3.500	700	250	1.200	8.000			13.650
#3	7	25.000	20.000	150.000	100	100	450	1.000	250.650
#4	3	1.000.000	400.000	220					1.400.220
#5	1	600							600

Tabella 13: esempio di simulazione Monte Carlo per il calcolo della distribuzione delle perdite aggregate (primi 5 scenari).

L'output d'interesse della simulazione è costituito dall'ultima colonna delle perdite totali in tabella 13; simulando un numero sufficientemente alto di scenari diversi e ordinando in senso crescente le perdite totali associate ad ognuno di essi si otterrà infatti la distribuzione aggregata cercata.

In questa ricerca sono state stimate, simulando 10.000 scenari, quattro diverse *aggregate loss distributions,* una per ognuno dei modelli *severity*/frequenza ritenuti adatti allo scopo tra quelli illustrati nelle pagine precedenti. In tabella 14 e 15 sono presentati i risultati della simulazione e le caratteristiche distributive dei modelli stimati; in particolare è possibile confrontare i percentili stimati per catturare le differenze tra questi modelli. I valori evidenziati in grassetto, contenuti nelle colonne relative ai percentili più alti (99% e 99,9%) non sono altro che i *Value-at-Risk* cercati. Nelle figure che seguono invece sono riportati gli istogrammi di ogni distribuzione aggregata simulata (i dati sono stati trasformati mediante logaritmo per una migliore visualizzazione).

Modello			**Percentili (in migliaia di *records*)**							
	Sev.	**Freq.**	**25%**	**50%**	**75%**	**90%**	**95%**	**99%**	**99.9%**	**Max**
(a)	LN	BN	0,22	2,9	33,7	305,4	1.193	**13.651**	**280.983**	1.089.656
(b)	LN + GPD (vincolo)	BN	0,21	2,5	26,4	209,2	882,6	**36.495**	**4.965.162**	3.101.124.462
(c)	LN + GPD (no vincolo)	BN	0,21	2,4	25,3	196,4	17.208	**77.952**	**220.343**	620.577
(d)	LN + GPD (no vincolo)	M-BN	0,16	2,1	23,5	191,2	18.110	**80.001**	**248.193**	689.366

Tabella 14: Distribuzioni aggregate delle perdite simulate per i 4 modelli scelti.

Modello	Statistiche descrittive (in migliaia)	
	Media	Dev. St.
(a)	1.256	21.522
(b)	681.007	37.102.567
(c)	3.348	19.316
(d)	3.417	19.705

Tabella 15: statistiche descrittive delle distribuzioni aggregate simulate.

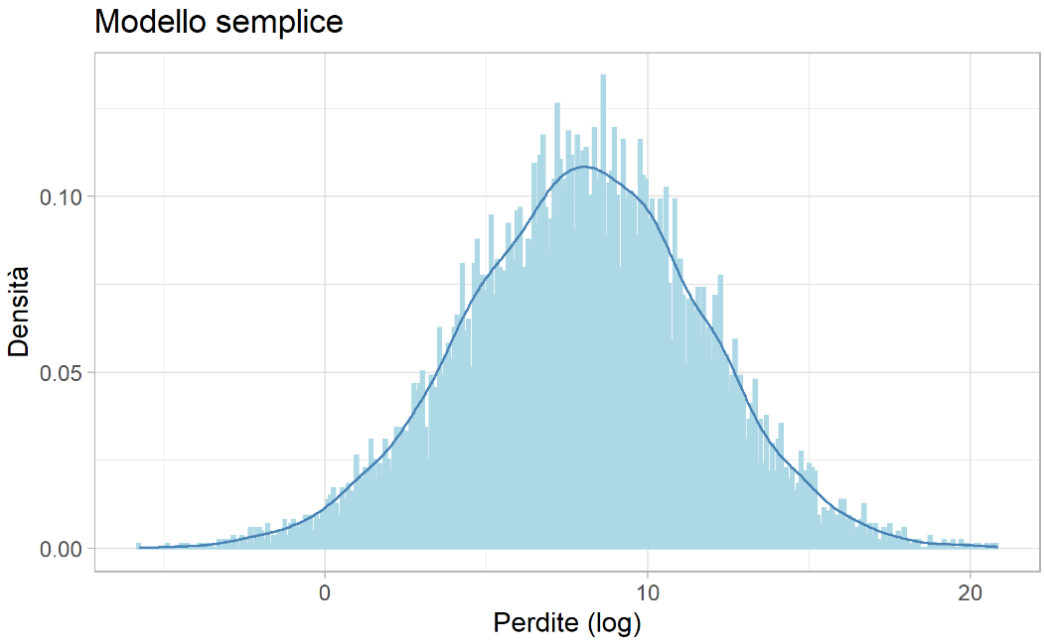

Figura 24: distribuzione aggregata delle perdite (in logaritmo) stimata attraverso il modello semplice (Log-Normale + Binomiale negativa).

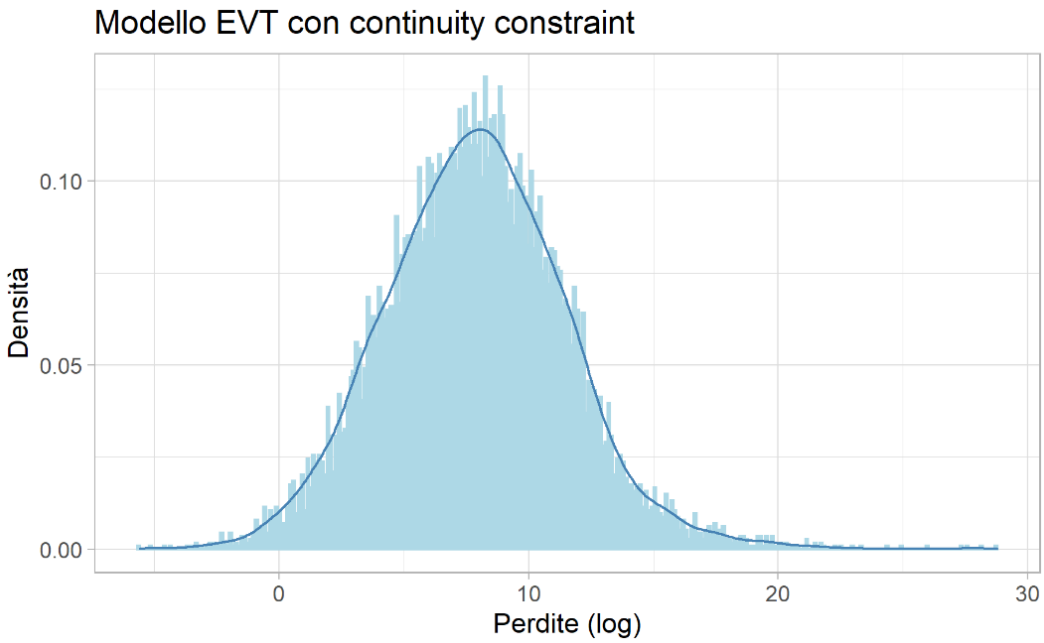

Figura 25: distribuzione aggregata delle perdite (in logaritmo) stimata attraverso il modello misto con vincolo di continuità tra corpo e coda della distribuzione.

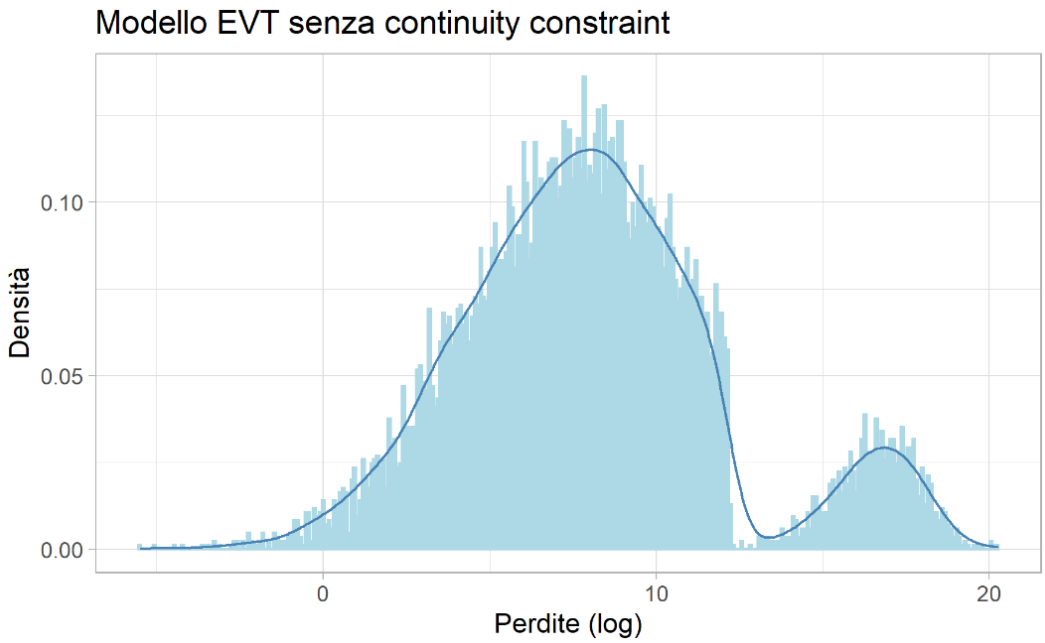

Figura 26: distribuzione aggregata delle perdite (in logaritmo) stimata attraverso il modello misto senza vincolo di continuità tra corpo e coda della distribuzione.

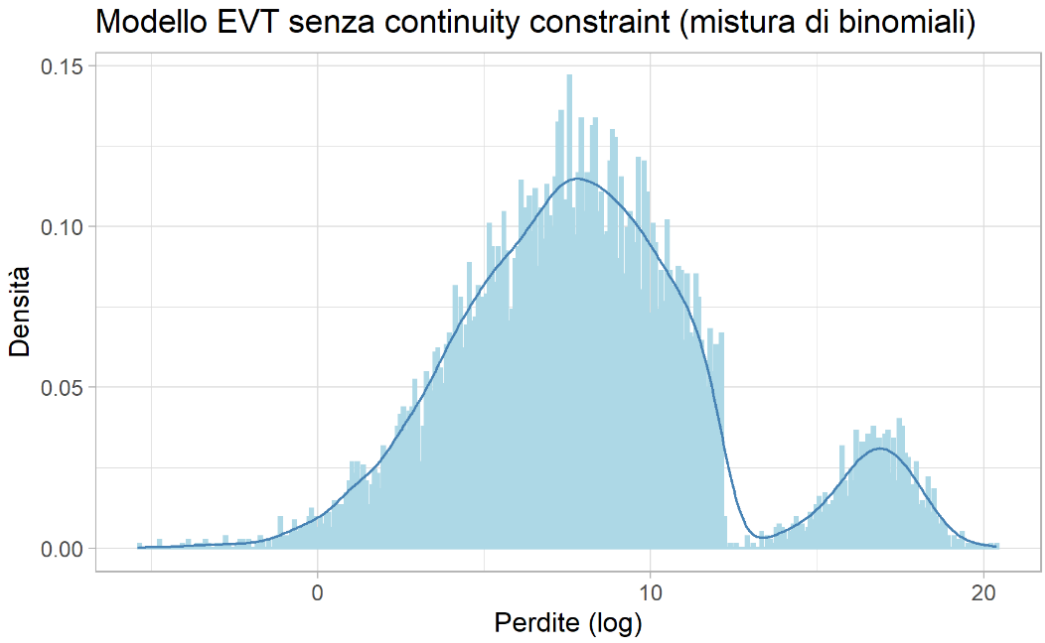

Figura 27: distribuzione aggregata delle perdite (in logaritmo) stimata attraverso il modello misto senza vincolo di continuità tra corpo e coda della distribuzione e frequenza descritta dalla mistura di binomiali negative.

3.4 Risultati e discussione

La prima cosa che risalta dall'osservazione delle tabelle 14 e 15 sono i valori assolutamente fuori scala, rispetto agli altri modelli, restituiti dal modello EVT con l'imposizione del vincolo di continuità, il quale presenta una coda decisamente più lunga. Il $VaR_{99,9\%}$ stimato tramite questo modello ammonta a circa 5 miliardi di *lost records*; lo stesso modello costruito senza vincolo di continuità restituisce invece un $VaR_{99,9\%}$ sensibilmente più basso (circa 220 milioni) e che sembra anche più coerente con i valori contenuti nel dataset. Su

questo punto, tuttavia, non è stato possibile elaborare una spiegazione teorica convincente data la scarsità di evidenze empiriche nella letteratura correlata sull'utilizzo di un simile vincolo di continuità. Seguendo una teoria derivata da Hu (2013), è possibile che il modello del "corpo" della distribuzione sia stato erroneamente specificato, nel qual caso l'imposizione del vincolo di continuità avrebbe un effetto dannoso sulla sensibilità del *fit* del modello nella coda. Tuttavia, la discontinuità tra corpo e coda evidenziata nel paragrafo 3.3.2 è, alla fine dei conti, di poca importanza in quanto incide solo sulla stima dei valori intorno alla soglia *u*; essendo l'oggetto d'interesse di questa ricerca costituito, in ultima analisi, dalla coda della distribuzione aggregata, e dati i risultati poco soddisfacenti e poco aderenti alla realtà (nonostante un *fit* statistico apparentemente migliore), decidiamo dunque di scartare il modello con vincolo di continuità.

Un'altra informazione interessante che si può desumere dalla tabella è la diversa qualità delle stime effettuate attraverso il modello *spliced* rispetto al modello semplice costituito dalla semplice distribuzione Log-Normale: quest'ultima infatti dovendosi adattare a tutto il range di dati a disposizione tende inevitabilmente a sovrastimare i dati nel "corpo" della distribuzione (percentili 25%, 50% e 75%) rispetto al modello *spliced*, mentre non mostra una coda sufficientemente pesante data la bassa probabilità associata agli eventi con perdite estreme. Questi problemi sono invece risolti in maniera naturale dal modello EVT, che consente di stimare adeguatamente tutto lo spettro dei possibili eventi abbassando anche la stima delle misure di rischio e, di conseguenza, il requisito di capitale (eventuale).

L'ultima considerazione riguarda il comportamento della distribuzione aggregata quando viene utilizzata come distribuzione di frequenza degli eventi una Binomiale Negativa semplice (d'ora in avanti SBN) o la mistura di Binomiali Negative (MBN) discussa nel paragrafo 3.3.1. Osservando i percentili delle due distribuzioni aggregate, si nota come il modello con MBN fornisca una stima leggermente più bassa nei percentili più bassi, cioè nel corpo, e più alta nei percentili oltre il 90%, cioè nella coda. Si tratta di un risultato che non sorprende in quanto era proprio quello ricercato, consentendo agli eventi "estremi" di verificarsi con una frequenza leggermente più alta rispetto a quella che una singola distribuzione avrebbe consentito, per rispecchiare meglio l'andamento empirico del fenomeno. Si ritiene dunque che entrambe le versioni del modello, SBN e MBN, siano sufficientemente adeguate e la scelta tra l'una e l'altra rappresenta il classico *trade-off* tra semplicità e accuratezza delle stime: da un lato, il principio di parsimonia favorirebbe il modello semplice; dall'altro, il calcolo della *mixture distribution* non rappresenta uno step particolarmente arduo e consente di ottenere, almeno apparentemente, una stima di miglior qualità.

In tabella 16 sono riportate le misure di rischio associate ad ognuno dei modelli stimati.

Modello	VaR 99%	ES 99%	VaR 99,9 %	ES 99,9%
Semplice (Log-Normale)	13.651.742	104.903.966	280.983.153	584.669.556
EVT Con vincolo	36.495.506	68.072.851.622	4.965.161.744	676.645.412.336
EVT Senza vincolo (SBN)	77.952.461	145.981.611	220.343.016	400.625.520
EVT Senza vincolo (MBN)	80.001.688	149.351.303	248.193.137	405.929.730

Tabella 16: misure di rischio stimate dai 4 modelli.

Si tratta, come già menzionato, di misure di rischio espresse in termini di *records* persi accidentalmente o sottratti in maniera fraudolenta. La conversione di queste quantità in unità di misura monetaria rappresenta un compito non semplice e necessariamente basato su stime approssimative. Un esempio chiarirà questo punto: è plausibile affermare che, superata una certa soglia di *record* sottratti, i costi medi associati alle spese legali, alla notifica ai clienti e/o alla risposta all'attacco vadano progressivamente a diminuire a causa dell'effetto dovuto ad economie di scala. Questo fatto è confermato dal report di IBM & Ponemon Institute[79], il quale stima un costo medio per record perso di $161 per i *data breach* con un numero di *record* persi pari o inferiore a 100.000; per gli eventi che invece coinvolgono un numero di *record* superiore al milione (i cosiddetti *mega breach*), il costo medio assume un andamento decrescente fino ad abbassarsi intorno ai $7 per *record* per gli eventi di maggior magnitudo.

Un possibile metodo di conversione è fornito da Jacobs (2014) e ripreso anche da Romanosky (2016) e Edwards, Hofmeyr and Forrest (2016). Egli mostra che il costo di un *data breach* può essere catturato dal seguente modello log-log:

$$\log(C) = 7{,}68 + 0{,}7584 \times \log(S) \qquad (3.3)$$

dove C rappresenta il costo dell'evento in dollari statunitensi mentre S rappresenta l'impatto (*size*) in termini di *record* persi. Applicando dunque questo metodo di conversione, otteniamo le misure di rischio (in termini monetari) mostrate in tabella 17.

Modello	($) VaR 99%	($) ES 99%	($) VaR 99,9 %	($) ES 99,9%
Semplice	558.095.158	2.620.295.093	5.531.727.099	9.642.874.831
EVT Senza vincolo (SBN)	2.091.921.409	3.366.552.097	4.600.320.846	7.239.333.487
EVT Senza vincolo (MBN)	2.133.496.903	3.425.324.640	5.034.892.256	7.311.908.586

Tabella 17: misure di rischio stimate in unità di misura monetaria (dollaro statunitense) ottenute attraverso il modello di Jacobs (2014).

[79] Ponemon Institute & IBM Security, Cost of a Data Breach Report (2021).

3.4.1 Applicazione del modello ad altri settori

L'analisi svolta in questo capitolo è stata fino ad ora incentrata sulla valutazione del *cyber risk* per il settore finanziario, che rappresenta l'oggetto principale d'interesse; può essere tuttavia interessante applicare in maniera comparativa la stessa metodologia a settori diversi per mostrarne le differenze nei profili di rischio e validare ulteriormente la tesi, fin qui sostenuta, di una maggiore esposizione complessiva del settore finanziario. Si è scelto pertanto di confrontare le stime per il settore finanziario con quelle del settore medico e del settore generico "business" che, come visto in apertura di Capitolo, risultano tra quelli maggiormente esposti sebbene con gradi di protezione diversi; a questo scopo, saranno confrontate le stime restituite dal modello "semplice" e quello EVT (SBN) opportunamente ricalibrati. È tuttavia necessario sottolineare che, per quanto riguarda le distribuzioni di frequenza e di *severity* sottostanti al modello, si è scelto di non modificarle, sebbene queste ultime non necessariamente costituiscano il miglior *fit* statistico possibile[80]. Le stime dei parametri per questi due settori insieme ai risultati dei test di significatività sono riportati in tabella 18.

Settore Sanitario						*GoF (p-value)*		
Distrib.	**Parametri stimati**					**K-S**	**A-D**	**C-vM**
Frequenza (BN)		μ = 9,616052 (0,04706071)	size = 6697777 (9,376311)				0,003	0,004
Severity (Log-Normale)		μ = 7,785225 (0,03477651)	σ = 2,174853 (0,02459068)			0	0,453	0,465
Severity (EVT)	μ = 7,744	σ = 2,079	β = 7,744	*u* = 29.157	ξ = 0,1831	0	0,342	0,35
Settore "Business"								
Frequenza (BN)		μ = 8,75184 (0,1168387)	size = 13,8952 (1,7627257)				0,02	0,01
Severity (Log-Normale)		μ = 9,183664 (0,1931713)	σ = 3,987012 (0,1365927)			0,06	0,24	0,64
Severity (EVT)	μ = 9,067	σ = 3,74	β = 387.195.245	*u* = 1.850.000	ξ = 0,2323	0,126	0,54	0,74

Tabella 18: stime dei parametri del modello e test di bontà di adattamento per gli altri settori considerati.

In tabella 19 invece sono riportate le misure di rischio calcolate (in termini di *records* persi o sottratti) per entrambi i settori; anche in questo caso può essere applicato il modello di conversione descritto dall'equazione (3.3).

[80] In particolare, nessuna delle distribuzioni di frequenza fin qui utilizzate sembra adattarsi in maniera soddisfacente ai dati, mentre la distribuzione Log-Normale e la distribuzione *spliced* rappresentano un ragionevole *fit* per quella di *severity*.

Settore Sanitario				
Modello	**VaR 99%**	**ES 99%**	**VaR 99,9 %**	**ES 99,9%**
Semplice	374.403	1.091.956	2.036.483	4.294.248
EVT Senza vincolo (SBN)	9.084.579	13.871.136	19.959.103	24.448.246
Settore "Business"				
Semplice	108.103.542	1.405.928.179	2.989.929.821	10.785.722.706
EVT Senza vincolo (SBN)	987.813.274	1.684.403.523	2.438.356.868	3.729.653.837

Tabella 19: misure di rischio stimate per i settori medico e business (non in termini monetari).

Si nota chiaramente come anche per questi sotto-campioni settoriali vi sia una notevole differenza tra le stime offerte dal modello "semplice" e dal modello EVT, a parità di altre condizioni; il modello semplice, infatti, sottostima notevolmente il 99° percentile della distribuzione aggregata e tende invece a dare una stima leggermente più alta del $VaR_{99,9\%}$, così come si era già osservato per il settore finanziario. Si evidenzia poi come il settore medico, sebbene sia il più colpito in termini assoluti tra i settori compresi nel dataset e anche il meno "protetto" (paragrafo 3.2), presenti un grado di rischiosità sensibilmente più basso se comparato con gli altri due settori considerati; al contrario, il settore business possiede il grado di rischiosità più alto, nonostante risulti anche quello meglio protetto. Il settore finanziario si va a porre esattamente a metà strada tra questi due settori, in termini di rischiosità. Se da un lato ciò sembra confutare la tesi, dall'altro è opportuno tenere a mente che il settore generico "business" è in questo caso un contenitore residuale, comprensivo di comparti economici molto diversi ed eterogenei tra loro: basti pensare che già i primi dieci eventi di maggior impatto in questo macrosettore riguardano entità appartenenti ad aree molto diverse, come fornitori di servizi online, società di marketing e *data broker*; chiaramente, aggregando i profili di rischio di questi settori così diversi tra loro non si otterrà una stima sufficientemente attendibile e confrontabile con quelle ricavate, al contrario, per il settore finanziario-assicurativo e quello sanitario.

3.5 Conclusioni, limitazioni e ricerca futura

Lo scopo di questo lavoro di tesi, pur con le limitazioni di seguito illustrate, è quello di fornire una panoramica qualitativa e quantitativa di un fenomeno che si ritiene acquisirà sempre più centralità in futuro e dalla cui adeguata gestione passerà la stabilità del sistema economico-finanziario, ma anche sociale, dei prossimi decenni. L'implementazione di modelli quantitativi per il *cyber risk* rappresenta un passo cruciale verso questa adeguatezza gestionale, in quanto permetterà, tra le altre cose: agli enti economici di indirizzare al meglio le proprie scelte in termini di investimento in *cybersecurity*; ai risparmiatori di scegliere gli istituti più virtuosi e

sicuri dove depositare i propri risparmi; agli enti assicurativi di fornire polizze su misura per questo rischio[81]; ai *policymaker* nazionali e sovranazionali di affinare sempre di più le politiche introdotte per combattere il fenomeno. Si tratta, indubbiamente, di una sfida non semplice in quanto il *cyber risk* ha conosciuto e continua a sperimentare una continua evoluzione.

I risultati ottenuti attraverso l'analisi svolta in questo lavoro sono, come già anticipato, da intendere come puramente illustrativi della dimensione e dell'importanza, anche in prospettiva futura, del fenomeno del *cyber risk*. È stata analizzata infatti solo una delle varie configurazioni in cui questo tipo di rischio può materializzarsi, e cioè i *data breach*, eventi di perdita accidentale o di furto di dati sensibili, che secondo Eling and Wirfs (2018) costituiscono solo il 25% del totale dei *cyber events*.

Un'analisi illustrativa di questo tipo, per sua natura, non può essere esente da limitazioni teoriche e pratiche; si è già diffusamente parlato, ad esempio, del problema relativo alla mancanza, almeno su base pubblicamente disponibile, di dati consistenti su cui poter effettuare analisi statisticamente robuste. Il dataset utilizzato in questo lavoro infatti è costituito solo da eventi pubblicamente riportati dai media e quindi per sua natura incompleto, ed è quindi probabile che calibrando il modello su di esso si vada a sottostimare la vera entità del fenomeno. Un'altra importante limitazione riguarda poi l'eterogeneità degli eventi inclusi nel nostro dataset, che riguardano organizzazioni finanziarie di diversa natura, grandezza (sia in termini di espansione geografica che di dimensione economica) e importanza; per questo motivo, i risultati in termini di rischio ottenuti vanno intesi, come già sottolineato, come complessivi per l'intero comparto e non a livello di singole organizzazioni. Infine, è opportuno sottolineare anche come si sia scelto di aggregare gli eventi "accidentali" e quelli "intenzionali" analizzandoli insieme come parte di un unico profilo di rischio. In questo caso, l'approccio più significativo consisterebbe nel trattare queste due differenti tipologie di eventi in maniera separata, in quanto diversi sono i rimedi e i controlli da attuare per ridurre il livello di rischio generale.

Alcune possibili espansioni della metodologia applicata possono essere: l'introduzione di una struttura di dipendenza tra eventi verificatisi in settori diversi per riprodurre un "effetto contagio" tra di essi; l'introduzione nel modello di un parametro temporale, che permetta ai parametri stimati di cambiare nel tempo per rappresentare meglio l'evoluzione dinamica del fenomeno; la suddivisione degli eventi su base geografica per analizzare anche l'impatto di differenti legislazioni e sistemi di sorveglianza sul livello complessivo di rischio[82];

[81] Infatti, come notato da Eling and Wirfs (2018) l'industria assicurativa sembra ancora riluttante ad offrire polizze contro il *cyber risk* date le difficoltà di quantificazione dello stesso.
[82] Il dataset qui utilizzato, si ricorda, è infatti limitato alle sole organizzazioni statunitensi.

l'implementazione di analisi di scenario e *stress testing* per valutare i possibili scenari futuri e, infine, la costruzione di un modello dei costi (diretti e indiretti) associati agli eventi che permetta di quantificare al meglio le perdite in termini monetari stimate attraverso il modello presentato.

Pur ferme le limitazioni di cui sopra, comunque, il lavoro presentato fa uso di una metodologia ben definita, di semplice comprensione, replicabile e adattabile, sperimentata con frequenza crescente negli anni dalla letteratura correlata e si ha dunque motivo di ritenere che con input di migliore qualità e ulteriore raffinamento, il modello creato attraverso tale metodologia possa in qualche modo rappresentare una valida alternativa.

Attraverso l'analisi svolta si è sottolineata l'importanza dell'applicazione degli strumenti della *Extreme Value Theory* per rappresentare adeguatamente un fenomeno in cui gli eventi di impatto estremo si verificano con una frequenza e una probabilità più alta di quanto i modelli probabilistici standard riuscirebbero a catturare; si è testato l'utilizzo di una distribuzione di frequenza "mista", che riuscisse a descrivere adeguatamente la frequenza delle due "modalità di eventi" presenti nel dataset, e si è infine accertata la significatività del modello anche per settori diversi da quello finanziario.

Bibliografia

[1] Aldasoro, I., Frost, J., Gambacorta, L., & Whyte, D. (2020). *Cyber risk in the financial sector*. *SUERF Policy Note*, *825215*(206), 1–15. www.bis.org

[2] Aldasoro, I., Frost, J., Gambacorta, L., & Whyte, D. (2021). *Covid-19 and cyber risk in the financial sector* (No. 37). Bank for International Settlements.

[3] Aldasoro, I., Gambacorta, L., Giudici, P., & Leach, T. (2020a). *Operational and Cyber Risks in the Financial Sector*. *Ssrn*, *8*.

[4] Aldasoro, I., Gambacorta, L., Giudici, P., & Leach, T. (2020b). *The Drivers of Cyber Risk*. In *BIS Working Papers* (Issue 865).

[5] Allen, L., Boudoukh, J. and Saunders, A. (2009) *Understanding Market, Credit, and Operational Risk*. 1st edn. Wiley. Available at: https://www.perlego.com/book/2788741/understanding-market-credit-and-operational-risk-pdf (Accessed: 25 September 2021).

[6] Allianz Global Corporate & Specialty (AGCS). (2021). *Allianz Risk Barometer: identifying the major business risks for 2021*.

[7] Artzner, P., Delbaen, F., Eber, J. M., & Heath, D. (1999). *Coherent measures of risk*. Mathematical Finance, 9(3), 203–228. https://doi.org/10.1111/1467-9965.00068

[8] AXA. (2021). *Future risks report (2021).*

[9] BCBS. (2021). Basel Committee on Banking Supervision, *Principles for Operational Resilience Principles for Operational Resilience III*. March. www.bis.org

[10] Bentley, M., Stephenson, A., Toscas, P., & Zhu, Z. (2020). *A multivariate model to quantify and mitigate cybersecurity risk*. Risks, 8 (2), 1–21. https://doi.org/10.3390/risks8020061

[11] Bouveret, A. (2018). *Cyber Risk for the Financial Sector: A Framework for Quantitative Assessment*. In IMF Working Papers (Vol. 18, Issue 143). https://doi.org/10.5089/9781484360750.001

[12] Buith J. & Spataru D. (2015), *The Benefits, Limits of Cyber Value-at-Risk*, The Wall Street Journal (online) & Deloitte. https://deloitte.wsj.com/articles/the-benefits-and-limits-of-cyber-value-at-risk-1432094595

[13] Calliess, C., & Baumgarten, A. (2020). *Cybersecurity in the EU the example of the financial sector: A legal perspective.* In German Law Journal (Vol. 21, Issue 6). https://doi.org/10.1017/glj.2020.67

[14] Cebula, J. J., Popeck, M. E., & Young, L. R. (2014). *A Taxonomy of Operational Cyber Security Risks Version 2*. In Carnegie-Mellon Univ Software Engineering Inst (Issue May, pp. 1–47). http://www.sei.cmu.edu

[15] Chernobai, A. S., Rachev, S. T., & Fabozzi, F. J. (n.d.). *Operational Risk A Guide to Basel II Capital Requirements, Models, and Analysis*. Retrieved February 13, 2022, from www.wiley.com.

[16] Cruz, M., Peters, G. and Shevchenko, P. (2015) *Fundamental Aspects of Operational Risk and Insurance Analytics*. 1st edn. Wiley. Available at: https://www.perlego.com/book/997213/fundamental-aspects-of-operational-risk-and-insurance-analytics-pdf (Accessed: 25 September 2021).

[17] Curti, F., Gerlach, J., Kazinnik, S., Lee, M. J., & Mihov, A. (2019). *Cyber risk definition and classification for financial risk management*. Federal Reserve Bank of St Louis, August, Mimeo.

[18] Edwards, B., Hofmeyr, S., & Forrest, S. (2016). *Hype and heavy tails: A closer look at data breaches*. Journal of Cybersecurity, 2(1), 3–14. https://doi.org/10.1093/cybsec/tyw003

[19] Eling, M., & Loperfido, N. (2017). *Data breaches: Goodness of fit, pricing, and risk measurement*. Insurance: Mathematics and Economics, 75, 126–136. https://doi.org/10.1016/j.insmatheco.2017.05.008

[20] Eling, M., Schnell, W., & Sommerrock, F. (2017). *Ten key questions on cyber risk and cyber risk insurance*. Asia Insurance Review, November, 92–93. http://search.ebscohost.com/login.aspx?direct=true&db=bsu&AN=121220137&site=ehost-live&scope=site

[21] Eling, M., & Wirfs, J. (2019). *What are the actual costs of cyber risk events?* European Journal of Operational Research, 272(3), 1109–1119. https://doi.org/10.1016/j.ejor.2018.07.021

[22] Eling, M., & Wirfs, J. H. (2015). *Modelling and Management of Cyber Risk.* International Actuarial Association, 1–20. https://www.actuaries.org/oslo2015/papers/IAALS-Wirfs&Eling.pdf

[23] ESRB. (2020). European Systemic Risk Board, *Systemic Cyber Risk*, February. https://www.esrb.europa.eu/

[24] Frachot, A., Georges, P. & Roncalli, T. (2001). *Loss Distribution Approach for Operational Risk.* Available at SSRN: https://ssrn.com/abstract=1032523 or http://dx.doi.org/10.2139/ssrn.1032523

[25] Frachot, A., Moudoulaud, O., Roncalli, T., & others. (2004). *Loss distribution approach in practice*. The Basel Handbook: A Guide for Financial Practitioners, 527–554.

[26] Financial Stability Board (FSB). (2012). *Cyber Lexicon* (Issue November).

[27] Gnedenko, B. (1943). *Sur La Distribution Limite Du Terme Maximum D'Une Serie Aleatoire* Author(s): B. Gnedenko Source: The Annals of Mathematics, Second Series, Vol. 44, No. 3, (Jul 1943), pp. 423-453 Published by: Annals of Mathematics Stable URL: http://. Annals of Mathematics, 44(3), 423–453. http://www.jstor.org/stable/1968974

[28] Greselin, F., Piacenza, F., & Zitikis, R. (2019). *Practice oriented and monte carlo based estimation of the value-at-risk for operational risk measurement*. Risks, 7(2). https://doi.org/10.3390/risks7020050

[29] Hu, Y. (2013). *Extreme Value Mixture Modelling with Simulation Study and Applications in Finance and Insurance*. Thesis, July. http://www.math.canterbury.ac.nz/~c.scarrott/evmix/thesis.pdf

[30] IBM. (2021). *Cost of a Data Breach Report 2021*.

[31] Kopp, E., Kaffenberger, L., & Wilson, C. (2017). *Cyber Risk, Market Failures, and Financial Stability*. IMF Working Papers, 17(185). https://doi.org/10.5089/9781484313787.001

[32] Krüger, P. S., & Brauchle, J.-P. (2021). *The European Union, Cybersecurity, and the Financial Sector: A Primer.* Cyber Policy Initiative Working Paper Series | "Cybersecurity and the Financial System," 9, 1–37.

[33] Li, J., Feng, J., & Chen, J. (2009). *A piecewise-defined severity distribution-based loss distribution approach to estimate operational risk: Evidence from chinese national commercial banks.* International Journal of Information Technology and Decision Making, 8(4), 727–747. https://doi.org/10.1142/S0219622009003727

[34] Maillart, T., & Sornette, D. (2010). *Heavy-tailed distribution of cyber-risks.* European Physical Journal B, 75(3), 357–364. https://doi.org/10.1140/epjb/e2010-00120-8

[35] McNeil, A. J. (1999). *Extreme Value Theory for Risk Manager: A General Introduction to Extreme Risk*. Internal Modelling and CAD II, 3, 1–22. http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Extreme+Value+Theory+for+Risk+Managers#0

[36] McNeil, A. J. (1997*). Estimating the Tails of Loss Severity Distributions Using Extreme Value Theory*. ASTIN Bulletin, 27(1), 117–137. https://doi.org/10.2143/ast.27.1.563210

[37] Moscadelli, M. (2004). *The Modelling of Operational Risk: Experience with the Analysis of the Data Collected by the Basel Committee*. Available at SSRN: https://ssrn.com/abstract=557214 or http://dx.doi.org/10.2139/ssrn.557214

[38] Nish, A., Naumaan, S., & Muir, J. (2020). *Enduring Cyber Threats and Emerging Challenges to the Financial Sector*. Carnegie Endowment for International Peace, November. https://carnegieendowment.org/2020/11/18/enduring-cyber-threats-and-emerging-challenges-to-financial-sector-pub-83239

[39] Orlando, A. (2021). *Cyber Risk Quantification: Investigating the Role of Cyber Value at Risk*. Risks, 9 (10), 184. https://doi.org/10.3390/risks9100184

[40] Panjer, H. (2006) *Operational Risk*. 1st edn. Wiley. Available at: https://www.perlego.com/book/2765550/operational-risk-pdf (Accessed: 12 January 2022).

[41] Parodi, P. (2014) *Pricing in General Insurance*. 1st edn. CRC Press. Available at: https://www.perlego.com/book/2193744/pricing-in-general-insurance-pdf (Accessed: 29 January 2022).

[42] Peters, G. and Shevchenko, P. (2015) *Advances in Heavy Tailed Risk Modeling*. 1st edn. Wiley. Available at: https://www.perlego.com/book/997045/advances-in-heavy-tailed-risk-modeling-pdf (Accessed: 8 January 2022).

[43] Sironi, A., & Resti, A. (2008). Rischio e valore nelle banche: misura, regolamentazione, gestione. *Rischio e valore nelle banche*, 1-955.

[44] Sophos. (2021). *The state of ransomware 2021*.

[45] Romanosky, S. (2016). *Examining the costs and causes of cyber incidents.* Journal of Cybersecurity, 2 (2), 121–135. https://doi.org/10.1093/cybsec/tyw001

[46] Shevchenko, P.V. (2010), *Implementing loss distribution approach for operational risk.* Appl. Stochastic Models Bus. Ind., 26: 277-307. https://doi.org/10.1002/asmb.812

[47] Strupczewski, G. (2021). *Defining cyber risk*. Safety Science, 135 (February 2020), 105143. https://doi.org/10.1016/j.ssci.2020.105143

[48] Verizon. (2021). *Data Breach Investigations Report.*

[49] WEF, World Economic Forum. (2015). *Partnering for Cyber Resilience: Towards the Quantification of Cyber Threats*. *January*, 1–20.

Riassunto

La progressiva digitalizzazione dell'economia globale a cui abbiamo assistito a partire dalla metà del secolo scorso ha fornito alle imprese nuove potenzialità e opportunità, ma ha enfatizzato anche il rischio collegato alla perdita della disponibilità o alla compromissione dell'integrità dei dati e/o dei sistemi di elaborazione delle informazioni, un rischio definito *Cyber Risk*. L'annuale "Future Risks Report" pubblicato da AXA, nel 2021, cita il cyber risk al secondo posto della Top 10 dei rischi emergenti per il prossimo decennio, dietro solo al rischio connesso al cambiamento climatico e, sorprendentemente, davanti al rischio portato da pandemie e malattie infettive.

La ricerca scientifica non ha ancora adottato uno standard comune di definizione del cyber risk, in quanto si tratta di una categoria di rischio complessa, che per essere trattata adeguatamente presuppone conoscenze sia tecniche che economiche e che presenta un forte grado di interdisciplinarità: contributi sull'argomento arrivano, tra gli altri, dai campi della finanza, delle assicurazioni, del risk management, ma anche dall'informatica e ovviamente dalla ricerca sulla cybersecurity. A titolo di esempio, l'Institute of Risk Management (IRM) definisce il cyber risk come "ogni rischio di perdita finanziaria, interruzione di servizio o danneggiamento della reputazione di un'organizzazione, derivante da un qualche tipo di guasto dei propri sistemi informatici"; ancora, sia il World Economic Forum (2012) che il Financial Stability Board nel Cyber Lexicon (2018) definiscono il cyber risk come "la combinazione tra la probabilità di realizzazione di un cyber incident e il relativo impatto", ove per cyber incident si intende "un evento che metta in pericolo la sicurezza di un sistema informatico o delle informazioni che il sistema processa, conserva o trasmette, oppure che violi le norme e le procedure di sicurezza, sia che tale evento derivi da attività malevola o meno". Non vi è un consenso generale su quali siano gli elementi fondamentali per definire il rischio IT, ed in effetti pare corretto affermare che "cyber risk" sia più un 'termine ombrello' sotto il quale è racchiusa una gamma di rischi diversi, risultanti da una disfunzione o da una violazione dei sistemi informatici.

Diversi autori in letteratura hanno sottolineato l'importanza di introdurre un sistema di classificazione degli incidenti cyber adeguato e standardizzato. Una prima categoria di eventi che è possibile individuare, come accennato in precedenza, è quella degli eventi accidentali. Questi possono avere origine interna o esterna all'impresa e sono il risultato di azioni che non prefigurano un intento malizioso, oppure semplicemente hanno origine naturale; danno luogo ad un rischio puramente operativo, idiosincratico, connaturato all'attività stessa d'impresa, ed

è dunque possibile prevenirli e gestirli adottando procedure e sistemi di controllo adeguati. Una seconda categoria, ben più vasta, è quella degli eventi intenzionali o man-made, cioè di quegli eventi che originano dall'intento di ledere all'organizzazione, e possono generarsi dall'interno (insider threat) o dall'esterno (cybercrime). I responsabili di questi attacchi (definiti threat actors nell'ambito della cybersecurity), le loro motivazioni, i mezzi e le strategie impiegate possono essere molteplici, e saperli riconoscere e distinguere può essere un importante vantaggio sia in fase di prevenzione che in fase di contenimento dei danni. Tra i metodi di attacco più diffusi possiamo citare: il malware, abbreviazione di "Malicious Software", termine che si riferisce ad una categoria di programmi ideati per danneggiare o trarre vantaggio da qualsiasi dispositivo programmabile, server o anche reti e usati dai cybercriminali principalmente per estorcere dati o informazioni da utilizzare come leva per conseguire un profitto; gli attacchi *Denial-of-Service* (DoS) o *Distributed-Denial-of-Service* (DDoS) che si verificano quando gli utenti legittimi non sono in grado di accedere ai sistemi informatici, ai dispositivi o alle risorse di rete a causa dell'azione di un *cyber threat actor;* il *phishing*, noto metodo di truffa attraverso il quale l'autore inganna le vittime inducendole a fornire dati sensibili, come password o dati bancari; o ancora gli attacchi "*man-in-the-middle"* in cui un terzo si inserisce clandestinamente nella comunicazione in corso tra due parti ignare, allo scopo di carpire informazioni e dati sensibili come credenziali di accesso o numeri di carte di credito.

Agli occhi delle autorità di regolazione, il *cyber risk* si configura come un (ampio) sottoinsieme della categoria dei rischi operativi sostenuti da un'impresa. Come il rischio operativo, infatti, anche il *cyber risk* è un rischio puro, non speculativo, che non dà luogo a possibili guadagni, ma solo a possibili perdite e che è legato a cause accidentali non prevedibili; parimenti, esso presenta una sorta di "bimodalità", ove ad incidenti molto frequenti sono associate piccole perdite, mentre ad eventi molto rari sono associate perdite di grande severità (*high frequency low impact vs low frequency high impact*). Se possibile, anzi, il cyber risk presenta ulteriori difficoltà di natura tecnica, come ad esempio nella costruzione di basi di dati complete ed esaustive in assenza di uno standard comune di rilevazione. Organizzazioni che sperimentano un cyber incident potrebbero essere restìe a divulgarlo per non ledere alla propria reputazione o potrebbero non avere i mezzi o le competenze tecniche necessarie a rilevarlo; per di più, anche qualora si riesca a costruire un dataset sufficientemente ampio e ben strutturato, le informazioni derivabili da esso potrebbero presto perdere di attendibilità: si tratta infatti di un ambito in rapida evoluzione a causa della incessante innovazione tecnologica.

Il cyber risk, tuttavia, presenta anche peculiarità difficilmente estendibili alla macrocategoria di riferimento: un attacco informatico mirato a una grande banca, ad esempio, ha potenzialmente una risonanza mediatica molto elevata, a cui si associano costi reputazionali

alti, ma che sono tradizionalmente esclusi dalla definizione di rischio operativo. In questo senso è esemplificativo lo studio di Eling and Wirfs (2018) che utilizzano un database di perdite 'operative' e ne estraggono quelle riconducibili a cyber events, mettendole a confronto con il resto del campione: i risultati dei test statistici effettuati dagli autori mostrano che i due sotto-campioni cyber e non-cyber sembrano provenire da popolazioni differenti, e anche il Value-at-Risk calcolato è più alto per le perdite cyber-related. Altro punto cruciale è che il cyber risk, specie se inteso come 'vulnerabilità' agli attacchi esterni, ha il potenziale per trasformarsi da rischio idiosincratico in rischio sistemico, attraverso noti effetti contagio. La vulnerabilità di una singola componente è infatti sufficiente per esporre tutto il sistema: esemplificativa è la nota vicenda del ransomware WannaCry che nel maggio 2017 arrivò ad infettare oltre 200.000 computer in 150 paesi, bloccando l'operatività di molte istituzioni pubbliche, banche, imprese e addirittura ospedali e richiedendo il pagamento di un riscatto per liberare i dati criptati dal virus.

Tra tutti, il sistema finanziario è uno dei principali canali attraverso cui il cyber risk può tramutarsi da rischio idiosincratico in rischio sistemico. Per la natura stessa dell'attività svolta, le banche e gli intermediari finanziari in generale risultano tra le imprese più esposte al cyber risk, in quanto fortemente dipendenti dal corretto funzionamento dei sistemi informatici e comunicativi e dei processi interni attraverso i quali si concretizzano materialmente le attività aziendali principali, come quella d'intermediazione o d'investimento; inoltre, detenendo grosse quantità di dati sensibili della propria clientela e movimentando imponenti volumi di denaro, gli istituti finanziari rappresentano un target troppo proficuo per i cybercriminali. Se da un lato, comunque, è ragionevole pensare che errori umani o disfunzioni accidentali non siano significativamente più frequenti per le imprese finanziarie piuttosto che per imprese di altri settori, lo stesso non si può dire per quanto riguarda le minacce provenienti dall'esterno, in particolare sottoforma di cybercrime o cyber terrorism. La frequenza degli attacchi informatici indirizzati a istituti finanziari, infatti, è da anni in costante crescita, come mostrato in Aldasoro et al (2020b).

Sebbene sia tra i più colpiti, comunque, il settore finanziario sembra sperimentare perdite relativamente più basse: circa $1.7 milioni per evento a fronte di una media di $2.6 milioni per tutti i settori; McKinsey (2021) nel resoconto di un sondaggio condotto presso più di 100 imprese e organizzazioni ha infatti riscontrato che il settore finanziario è tra i primi 3 per livello medio di sviluppo della cybersecurity; si tratta di un risultato che in effetti non sorprende più di tanto, in quanto si tratta di un settore che da un lato dipende fortemente dalla fiducia dei consumatori dovendone trattare e conservare i dati sensibili oltre che i risparmi, e che dall'altro si trova sotto la lente d'ingrandimento delle autorità di regolazione e di controllo.

La legislazione europea in materia di cybersecurity per il settore finanziario è abbastanza complessa e strutturata su più livelli; non esiste una legislazione unica, quanto più una moltitudine di atti (leggi, regolamenti, direttive, linee guida) che non in tutti i casi hanno valore legalmente vincolante. Ad un primo livello generale si pongono due direttive di ampio respiro, non specifiche di settore, emanate nel 2016: la Directive on Security of Network and Information Systems (Direttiva NIS) e il General Data Protection Regulation (GDPR). La direttiva NIS, recepita in Italia con il d.lgs. 18 maggio 2018, n.65 e attualmente in fase di revisione da parte dell'Unione, mira a definire le misure necessarie per ottenere un adeguato livello di sicurezza delle reti e dei sistemi informatici; il GDPR, che a differenza della NIS è operante a livello comunitario senza bisogno di ricezione nei singoli ordinamenti nazionali, mira poi a fornire una legislazione standard unica in materia di protezione dei dati e a dare un maggiore controllo ai cittadini su come i propri dati vengono utilizzati; in questo senso, si applica a tutti quegli enti che processano o controllano dati personali e che operano nel territorio dell'Unione.

A queste fonti normative di ampio respiro si affianca poi un corpus di leggi, direttive e regolamenti specifici per ogni comparto del settore finanziario; ad esempio, per gli istituti di credito l'insieme di norme esistente è particolarmente complesso e stratificato, in quanto trattasi di organizzazioni dall'elevato grado di digitalizzazione, di rilevante importanza sistemica e che offrono diversi servizi oltre a quelli legati alla tradizionale attività bancaria, in particolare servizi di pagamento e d'investimento. Un primo insieme di disposizioni in materia proviene dal pacchetto CRR/CRD IV, provvedimenti attuativi del Terzo Accordo di Basilea (Basilea III) nell'ambito dell'Unione in cui, riflettendo la visione del Comitato, la regolamentazione del cyber risk è trattata in maniera implicita in quanto considerato sottoinsieme del rischio operativo. Tra gli altri provvedimenti rilevanti per gli istituti di credito si rilevano la direttiva PSD2, rivolta ai fornitori di servizi di pagamento, e la MIFID II, direttiva rivolta ai fornitori di servizi d'investimento. È importante citare anche il Cyber Incident Reporting Framework che si applica a tutte le banche di rilevanza sistemica, poste direttamente sotto l'egida della BCE; le disposizioni contenute nel framework pongono l'obbligo, in capo a queste banche, di riportare direttamente alla BCE i cyber incident di maggior rilevanza.

In questo lavoro si è scelto di approcciare all'analisi da una prospettiva di puro risk management, tralasciando gli aspetti strettamente tecnici in materia di cybersecurity e gli aspetti organizzativi relativi, ad esempio, ai processi decisionali e alle attribuzioni di responsabilità aziendale. Lo scopo dell'analisi sarà infatti quello di fornire una singola misura, quella del

Value-at-Risk relativo al rischio informatico, e di discuterne i potenziali usi. A questo fine si è scelto di utilizzare gli strumenti adottati comunemente in materia di quantificazione del rischio operativo, e in particolare si farà riferimento al metodo denominato Loss Distribution Approach (LDA), tipico delle scienze attuariali e già utilizzato da diversi autori per scopi simili. In estrema sintesi, il LDA consiste nel trovare e combinare la distribuzione di frequenza degli eventi che causano perdite con la distribuzione di severità di queste ultime; a partire da questa combinazione, attraverso tecniche di simulazione come il metodo Montecarlo, viene generato un numero predeterminato di scenari sulla base dei quali si otterrà la distribuzione di probabilità aggregata delle perdite in un dato orizzonte temporale, dalla quale a sua volta sarà possibile ricavare gli indicatori di rischio cercati.

Un indicatore puramente finanziario di questo tipo può aiutare il personale non-tecnico a comprendere l'importanza di porre in essere comportamenti e processi idonei a preservare la sicurezza dell'organizzazione; può essere utile ad orientare le scelte d'investimento in sicurezza informatica da parte del management; e in ultima analisi garantire che adeguate riserve di capitale siano poste a salvaguardia dell'azienda per proteggerla da una minaccia non ancora pienamente compresa, in costante evoluzione e crescita e che rischia di rappresentare una zavorra per lo sviluppo futuro dell'intero settore.

Sviluppato a partire dalla prima metà degli anni Ottanta per la misurazione del rischio di mercato, il concetto di VaR ha da allora progressivamente acquisito importanza fino a divenire il modello prevalente nel campo della misurazione e della gestione dei rischi in generale, fino ad esser preso come riferimento anche dalle autorità di regolazione nel 1992 col primo accordo di Basilea. Gran parte di questa importanza è dovuta soprattutto alla semplicità concettuale del VaR, il quale essenzialmente punta a fornire la risposta alla domanda:

> *«Qual è la perdita massima a cui si può andare incontro in un determinato periodo temporale, tale che vi sia una probabilità molto bassa che la perdita effettiva risulti superiore al valore stimato?»*

I modelli VaR costituiscono una famiglia di tecniche diverse accomunate da questo impianto concettuale di base. Per quanto riguarda la materia in esame, e cioè la quantificazione del cyber risk, i due approcci principali sono quello basato sulla distribuzione parametrica e quello basato sulla distribuzione empirica delle perdite. In breve, nel primo approccio si cerca di modellare i dati a disposizione sulla base di una distribuzione conosciuta (ad esempio la distribuzione esponenziale o la Weibull) attraverso metodi di stima; nel secondo approccio (detto anche della simulazione storica) invece non si fa alcuna assunzione riguardo la forma funzionale della distribuzione, e si procede direttamente a calcolare il VaR sulla distribuzione

empirica delle perdite ricavata dai dati a disposizione. Il Loss Distribution Approach seguito nell'analisi presente in questo elaborato rientra nel primo caso. Un'alternativa al VaR è poi costituita dall' Expected Shortfall, una misura di rischio strettamente collegata al VaR che però ne supera alcune limitazioni teoriche. Esso risponde alla domanda:

«Se la perdita reale superasse il valore stimato dal VaR,
a quanto ammonterebbe la perdita attesa?»

La prima proposta organica per l'applicazione di una misura di rischio tradizionalmente finanziaria come il VaR nell'ambito della sicurezza informatica fu formulata nel 2015 nell'ambito dell'iniziativa per la Cyber-Resilience patrocinata dal World Economic Forum, che diede vita al framework concettuale denominato Cyber Value-at-Risk (Cy-VaR). Sebbene non fornisca un vero e proprio approccio operativo standardizzato per la quantificazione del cyber risk, la proposta del WEF contiene tuttavia delle indicazioni preziose ai fini della costruzione di un metodo completo e trasversale ai diversi settori dell'economia, incoraggiando poi le singole imprese a costruire i propri modelli di quantificazione interni. Il modello concettuale del Cy-VaR richiede alle imprese, innanzitutto, di comprendere sia i fattori chiave necessari alla costruzione di un modello per il cyber risk che le dipendenze e le interazioni esistenti fra di essi; le componenti fondamentali sono essenzialmente tre: la vulnerabilità dei sistemi aziendali, gli asset che potenzialmente possono essere obiettivo di attacchi informatici e il profilo del threat actor. Il Cy-VaR è dunque una misura di tipo economico che si presta ad essere utilizzata come complemento alle valutazioni strettamente tecniche sul livello di sicurezza informatica di un'impresa, e che quantificando l'impatto economico degli attacchi cyber aiuta anche ad orientare le scelte dei manager riguardanti gli investimenti in materia di cybersecurity e mitigazione del rischio (ad esempio, l'acquisto di una polizza assicurativa) o per quantificare la riduzione dell'esposizione derivante da queste scelte.

La metodologia utilizzata in questo lavoro, il Loss Distribution Approach, rientra tra i metodi suggeriti dal comitato di Basilea per la gestione del rischio operativo; in particolare, esso rientra tra gli Advanced Measurement Approaches, una famiglia di approcci di tipo bottom-up, che consistono nell'analisi del rischio operativo partendo "dal basso", ossia partendo dai dati raccolti all'interno della banca stessa, mappando ogni risk event in categorie specifiche e solo successivamente aggregando i dati per costruire, ad esempio, analisi di scenario o stress test. Si tratta dunque di approcci che incorporano il non banale vantaggio di spiegare i meccanismi che portano una banca ad avere una determinata esposizione di rischio, e che sono per questo motivo anche più complessi da implementare. Questa metodologia si è

nel tempo affermata come lo standard di riferimento nell'ambito della modellizzazione del rischio operativo. Come si legge in un documento consultivo a supporto dell'accordo di Basilea II, una banca che utilizzi il LDA deve stimare una funzione di distribuzione di probabilità per la severità e per la frequenza di ogni combinazione (o "cella") tra business line/event type su un orizzonte di un anno, combinando poi le due distribuzioni per ottenere una funzione di distribuzione cumulata per le perdite operative per ogni combinazione. Generalmente parlando, è improbabile che si ottenga una forma analitica nota e trattabile per la distribuzione cumulata delle perdite Zj, per cui è necessario ricorrere ad algoritmi numerici, come il metodo MonteCarlo, per ottenerne un'approssimazione. Aggregando le distribuzioni cumulate delle varie celle, si ottiene la distribuzione cumulata delle perdite operative totali per la banca; una volta ottenuta la distribuzione cumulata delle perdite, il passo successivo consiste nel calcolo degli indicatori di rischio.

Si tratta di un approccio non esente da limitazioni, e nel tentativo di superarle diverse soluzioni e metodi differenti sono stati proposti in letteratura. Un primo problema è relativo alla ben nota scarsità di dati utilizzabili per la stima delle distribuzioni di frequenza e severità delle perdite, in quanto a differenza del rischio di mercato o di credito, la raccolta di dati per il rischio operativo è iniziata in tempi relativamente più recenti. Per superare questo importante ostacolo, le soluzioni più diffuse consistono nell'utilizzo di distribuzioni parametriche a cui adattare i dati, e l'aggregazione di dati interni e di dati esterni (dunque provenienti da fonti diverse) per espandere i dataset disponibili. Nel primo caso si procede dunque a cercare di adattare i dati disponibili ad alcune distribuzioni di probabilità note, stimandone i parametri e la relativa incertezza di stima, verificando poi la bontà di adattamento del modello. L'aggregazione di dati interni e di dati esterni è invece un requisito richiesto in Basilea II per la qualificazione dei modelli AMA.

Un secondo problema, strettamente collegato al primo, riguarda invece la natura dei dati sulle perdite operative. È infatti noto che, molto più di altre tipologie di rischio, il rischio operativo si caratterizza per perdite di notevole entità che però occorrono con frequenza molto bassa, contrapposte a perdite di severità contenuta che occorrono con frequenza più alta. Ciò comporta inevitabilmente una difficoltà ulteriore nella costruzione di modelli adeguati, in quanto i dati riguardanti le perdite più severe potrebbero essere pochi o addirittura assenti nei dataset a disposizione. Per questo motivo, nella modellizzazione della severity delle perdite, una delle soluzioni più comunemente adottate (oltre all'utilizzo di dati interni ed esterni) comporta l'utilizzo di una distribuzione giunta (spliced distribution) ove il "corpo" della distribuzione, relativo alle perdite più frequenti e per cui sono quindi disponibili più dati, è

modellato secondo una distribuzione comune (ad esempio la Weibull, la Log-Normale o anche la distribuzione empirica), mentre la coda della distribuzione, relativa alle perdite più rare ma di maggior impatto è tipicamente caratterizzata attraverso tecniche della Extreme Value Theory.

Il primo passo per l'implementazione del LDA consiste nella stima della distribuzione di frequenza delle perdite; tra le distribuzioni parametriche più utilizzate per la stima della frequenza delle perdite troviamo la distribuzione di Poisson, senza dubbio la scelta principale e più diffusa nella letteratura sul rischio operativo e in quella più recente sul cyber risk. Si tratta di una distribuzione di probabilità discreta, utilizzata per trovare la probabilità del verificarsi di un certo numero k di eventi in un dato intervallo temporale. Essa presenta alcuni notevoli vantaggi. Il primo consiste nella dipendenza da un unico parametro, λ (chiamato anche intensity rate), che ne rappresenta sia il valore atteso che la varianza, a favore dunque di una maggiore semplicità; un altro importante vantaggio è dato dal fatto che, supponendo che X e Y siano due v.c. indipendenti e descritte entrambe da una Poisson con parametri λ_X e λ_Y, allora la distribuzione di X + Y è a sua volta una Poisson con parametro $\lambda_X+\lambda_Y$; si tratta di una proprietà utile quando, ad esempio, si voglia valutare la frequenza delle perdite da due o più business line. Infine, come evidenziato da Panjer (2006), se la frequenza delle perdite si distribuisce secondo una Poisson, e le perdite possono essere classificate in un numero n di categorie, seguirà che anche le perdite in ogni categoria seguono delle Poisson con parametri λ diversi.

Un potenziale problema derivante dall'utilizzo di una distribuzione di Poisson deriva dalla natura del parametro λ, che è assunto essere costante. Si tratta evidentemente di un'assunzione poco realistica, in quanto è probabile che esso vari nel tempo o che abbia un comportamento randomico. Per rimuovere questa assunzione poco realistica, una possibile soluzione consiste nell'assumere che λ segua un processo stocastico, e che dunque si evolva nel tempo secondo una funzione matematica $\lambda(t)$ (processo di Poisson non-omogeneo con intensità stocastica); oppure si può assumere che λ segua una distribuzione di probabilità a sua volta, ottenendo una mistura di distribuzioni (mixture distribution).

Un caso speciale di mistura di distribuzioni è la distribuzione Binomiale Negativa (BN), anch'essa molto utilizzata nella letteratura correlata; infatti, se si assume che il parametro λ della Poisson segua a sua volta una distribuzione Gamma, si ottiene proprio tale distribuzione. A differenza della Poisson, la Binomiale Negativa è una distribuzione tipicamente descritta da due parametri e che dunque ammette una maggiore flessibilità nella forma. Intuitivamente, essa rappresenta il numero di "fallimenti" che incorrono in una serie di prove di Bernoulli prima di ottenere un determinato numero di successi.

La stima della distribuzione di probabilità della dimensione delle perdite è, invece, un passaggio un po' più delicato. Infatti, mentre nella stima della frequenza è ragionevole e non troppo costoso – in termini di accuratezza delle stime – fare delle assunzioni semplificatrici (qual è l'adozione della Poisson, ad esempio), lo stesso non vale per questa fase del processo. Tra le distribuzioni di severity più diffusamente utilizzate in letteratura ricordiamo la distribuzione esponenziale, la distribuzione Log-Normale, di Weibull, e la distribuzione di Pareto; si tratta di distribuzioni heavy-tailed, ad eccezione dell'esponenziale, e che si prestano dunque bene a riflettere la tipica natura dei dati sul rischio operativo e, similmente, sul rischio cyber.

Tuttavia, data la già menzionata "bimodalità" di queste categorie di rischio, molti autori ricorrono alla Extreme Value Theory per poter caratterizzare adeguatamente ogni parte della distribuzione di severity. La Extreme Value Theory (EVT) è una branca della statistica che si occupa, come intuibile dal nome, dello studio dei valori estremi delle distribuzioni di probabilità – ossia quei valori che si allontanano fortemente dalla porzione centrale delle stesse (il "corpo"). Le tecniche della EVT sono spesso utilizzate nello studio dei rischi, e in particolare nello studio del rischio operativo, in quanto conferiscono la possibilità di valutare le code di una distribuzione anche in presenza di dati limitati. In altre parole, la EVT si occupa dello studio degli eventi più rari, caratterizzati da una bassa probabilità di verificarsi, ma che possono avere conseguenze potenzialmente catastrofiche. Per questo motivo, si tratta di una teoria particolarmente attrattiva per coloro che operano nel campo della gestione dei rischi, spesso chiamati a tenere conto di questo tipo di eventi nelle loro valutazioni. La EVT si fonda essenzialmente due tipi di modelli, o approcci, fondamentali: l'approccio Block Maxima (BM) e il relativamente più moderno approccio Peaks-over-Treshold (POT).

Nell'approccio Block Maxima si procede essenzialmente alla suddivisione dei dati in "blocchi" temporali di uguale ampiezza, e da ogni blocco si estrae la singola osservazione di maggior entità (ossia il massimo); si dimostra che, per valori sufficientemente alti, la distribuzione di tali massimi "normalizzati" estratti dai blocchi converge alla Generalized Extreme Value Distribution (GEV).

L'approccio POT consiste nello studio dei valori che superano una determinata soglia u, fissata ad un livello sufficientemente alto. Uno dei risultati chiave della EVT, che va sotto il nome di Teorema di Pickands-Balkema-de Haan, consiste nel fatto che, per un'ampia gamma di distribuzioni, la distribuzione dei dati al di sopra di questa soglia converge ad una distribuzione di Pareto generalizzata (GPD) all'aumentare della soglia u. Grazie a questo risultato, è possibile dunque considerare la GPD come la scelta più ovvia per andare a modellare la distribuzione delle perdite in eccesso oltre una certa soglia. La scelta del valore soglia u

rappresenta uno dei passaggi chiave dell'applicazione del metodo; come evidenziato da McNeill (1999), bisogna essenzialmente trovare un compromesso tra la scelta di un valore sufficientemente alto, tale da permettere l'applicazione delle proprietà asintotiche appena descritte, e la scelta di un valore abbastanza basso da avere dati sufficienti per la stima dei parametri della distribuzione. Si tratta di un problema per il quale tuttavia non vi sono ancora soluzioni universalmente accettate nella pratica: un possibile approccio, frequentemente utilizzato, si basa sull'analisi visiva del grafico della mean excess function (funzione dell'eccesso medio), definita come la media di tutte le differenze tra i valori che superano la soglia u e u, per diversi valori della soglia.

La scelta delle distribuzioni che meglio rappresentano la frequenza e l'onerosità delle perdite costituisce, senza dubbio, la parte più delicata e complessa della costruzione di un simile modello di quantificazione dei rischi. È chiaro che non è possibile ricavare un modello inequivocabilmente superiore agli altri, ma che piuttosto bisogna puntare a costruire un modello sufficientemente accurato sulla base degli strumenti e dei dati che si hanno a disposizione e attraverso un meccanismo che possiamo definire di trial-and-error; un processo di questo tipo passa, necessariamente, attraverso la conferma o la smentita delle ipotesi fatte a monte. Il primo passo nella conduzione di questo processo consiste nell'analisi grafica: attraverso l'osservazione di alcuni particolari grafici è infatti possibile iniziare a cogliere informazioni utili per il prosieguo dello studio, informazioni che poi dovranno essere confermate o smentite attraverso metodi numerici più formali; a dispetto di ciò, tale metodo rappresenta un ottimo punto di partenza in quasi ogni circostanza. Il primo dei grafici che generalmente si va ad osservare è quello della funzione di distribuzione cumulata empirica (ECDF), detta anche funzione di ripartizione empirica; un'ulteriore possibilità consiste nel confrontare l'istogramma dei dati con la densità della distribuzione stimata; tuttavia, il ricorso a tale grafico può risultare più problematico e meno immediato in alcuni casi, in quanto potrebbero occorrere problemi nel raggruppamento dei dati per generare l'istogramma. Ad ogni modo, entrambi i metodi risultano particolarmente sensibili alla presenza di outliers che rischiano di distorcere le rappresentazioni grafiche.

Un altro tipo di rappresentazione tipicamente molto utile è il Q-Q plot; in un grafico di questo tipo, i quantili della distribuzione empirica vengono confrontati con i quantili della distribuzione teorica ipotizzata. Intuitivamente, il fit fra le due distribuzioni sarà tanto migliore quanto più vicini saranno i punti del grafico ad una retta inclinata a 45°. L'ultimo tipo di grafico che vale la pena menzionare in questa sede è il Mean Excess Plot. Si tratta di un grafico spesso utilizzato in situazioni in cui si cerca di adattare un modello heavy-tailed ai dati, come spesso

accade nell'analisi dei rischi operativi e cyber; esso si rivela molto utile, ad esempio, come "guida operativa" per la scelta della soglia nell'applicazione del metodo POT; a seconda del trend che assumono i punti del grafico è possibile determinare che tipo di distribuzione si sta osservando: un'inclinazione verso l'alto suggerisce che si è in presenza di una distribuzione heavy tailed, mentre un'inclinazione verso il basso indica l'esatto opposto; una linea orizzontale, invece, suggerisce che si è in presenza di una distribuzione esponenziale. Un'altra informazione molto utile che è possibile ottenere riguarda il valore ottimale di u: se il grafico assume un'inclinazione positiva oltre un certo valore soglia, è possibile inferire che oltre quel valore i dati si distribuiscano secondo una GPD.

Come affermato in precedenza, le informazioni che è possibile desumere attraverso l'analisi di grafici come quelli appena descritti non sempre si rivelano corrette, o più semplicemente necessitano di validazione. Per accertare dunque la qualità del modello teorico ipotizzato è dunque necessario ricorrere a tecniche numeriche più formali quali i test d'ipotesi. Tipicamente, nell'ambito del framework delineato in questo capitolo, l'ipotesi nulla e alternativa che si vanno a testare sono le seguenti:

H0: i dati seguono la distribuzione ipotizzata

vs

H1: i dati non seguono la distribuzione ipotizzata

I test utilizzati in questa analisi sono i tre più comunemente utilizzati anche nella letteratura correlata: il test di Kolmogorov-Smirnov, il test di Anderson-Darling e quello di Cramér-von Mises. Si tratta di test basati sull'analisi della distanza tra la CDF teorica stimata e quella empirica, utilizzando però statistiche test differenti. Come regola operativa si guarda in genere al p-value fornito da tutti i software statistici: se tale valore è inferiore a 0.05, si rifiuta l'ipotesi nulla.

Il test di KS è sicuramente tra i più popolari, ma non è esente da difetti. Il principale svantaggio di questo tipo di test è dovuto al fatto che tende a sovrastimare il peso dei quantili più vicini alla mediana e di conseguenza a sottostimare il peso di quelli più lontani; in altre parole, il peso delle code verrà sottostimato dal test, fatto che rende il test poco ideale qualora si vada a testare il fit di modelli heavy-tailed; il test di Anderson Darling invece pone maggior peso sul fit delle code piuttosto che del corpo della distribuzione, ed è quindi maggiormente indicato quando si ritiene che i dati a disposizione siano heavy-tailed.

I risultati dell'analisi condotta attraverso la metodologia del Loss Distribution Approach sono presentati nel capitolo 3. Tale analisi è stata svolta partendo da uno dei dataset pubblici

più diffusi in materia, il "Data Breach Chronology" (DBC) costruito e reso disponibile dalla Privacy Rights Clearinghouse (PRC), un'organizzazione no-profit statunitense che si occupa di protezione dei consumatori e della loro privacy. Si tratta di un dataset abbastanza ampio, contenente informazioni su oltre novemila data breach avvenuti tra il 2005 e il 2019 a discapito di organizzazioni americane; sono escluse tuttavia dal dataset tutte quelle fattispecie di diversa natura (descritte nel Capitolo 1) che vanno a costituire l'insieme dei cyber risk. Nonostante ciò, i data breach rappresentano sicuramente una delle maggiori componenti di tale insieme, e pertanto si ritiene che l'analisi effettuata sia sufficientemente rappresentativa del fenomeno. Il dataset Data Breach Chronology contiene informazioni su 9015 data breach subiti da organizzazioni statunitensi, pubbliche o private, appartenenti a diversi settori e in un arco di tempo che va dal 2005 al 2019. Le informazioni fornite dal DBC utili ai fini della nostra analisi sono essenzialmente 4: anno dell'evento; numero di personal records rubati nell'evento; tipo di organizzazione (macrosettore di appartenenza); tipo di attacco.

Nel dataset, il settore in assoluto più colpito è quello sanitario, seguito dal macrosettore generico 'BSO' e dal settore finanziario. Si tratta di dati che in effetti confermano la rilevanza anche a livello "sistemico" del cyber risk, se si pensa alla centralità di cui il settore sanitario e quello finanziario godono nella vita sociale di ogni Stato moderno. Un'informazione fondamentale fornita dal dataset riguarda il numero di records persi o illegalmente sottratti da terzi per ogni attacco. Innanzitutto, è importante sottolineare come ben 2187 degli eventi registrati nel dataset (poco più del 24% delle osservazioni) non abbiano portato conseguenze, con nessun record perso; a tal proposito, si può pensare alla percentuale di eventi "nulli" come una proxy del grado di protezione di ogni macrosettore incluso nel dataset. I settori meglio protetti risultano il settore generico-industriale (addirittura quasi 60% di eventi che non hanno comportato perdita di dati), quello retail (poco più del 50%) e il settore finanziario-assicurativo (poco meno del 50%). Quest'ultimo dato in particolare è rappresentativo della differenza nel grado di protezione, e di riflesso di investimenti in cybersecurity, tra il settore privato e quello pubblico; una differenza che diventa allarmante considerando che il settore sanitario, l'altro settore di rilevanza sistemica oltre a quello finanziario, risulta drammaticamente esposto in quanto riesce a "neutralizzare" solo il 10% degli incidenti su un totale di 4343 osservazioni.

La restante parte del dataset è costituita poi dagli eventi che hanno effettivamente portato ad una perdita di dati sensibili. I dati inerenti a questi eventi sono caratterizzati da una forte asimmetricità e variabilità, come confermato anche dai valori di deviazione standard: sia per il dataset completo che per i tre sottoinsiemi, la differenza tra terzo quartile e valore massimo è ampissima. È questa un'informazione che conferma quanto affermato nel Capitolo 1 riguardo la "bi-modalità" di questo tipo di rischio: il corpo centrale dei dataset è formato dagli

eventi più frequenti ma a basso impatto, ma sono poi presenti anche eventi molto più rari ma dall'impatto molto più devastante.

Applicando il LDA è stato possibile dunque creare un modello teorico per la caratterizzazione del cyber risk; in realtà, poiché tale modello è stato calibrato sul dataset DBC, esso va a descrivere, più precisamente, un sottoinsieme di questa categoria di rischio, ossia il rischio collegato ai data breach. Inoltre, la calibrazione iniziale dei parametri e la scelta delle distribuzioni più adatte a descrivere i dati del modello è stata effettuata sul sottoinsieme del dataset comprendente i soli eventi che hanno colpito organizzazioni classificate nel settore finanziario. Nonostante ciò, si tratta di un modello che può essere facilmente esteso e ricalibrato. Per la distribuzione di frequenza, si è ipotizzata una distribuzione di Poisson o una Binomiale Negativa; in particolare, con entrambe le tipologie di modelli si è cercato anche di costruire una distribuzione "mista", costituita da due distribuzioni l'una riferita agli eventi ad alta frequenza ma a basso impatto e l'altra riferita agli eventi rari ma di maggior impatto. Poiché tutti i test d'ipotesi ci portano a rifiutare la Poisson come distribuzione adeguata, tuttavia, questa è stata completamente scartata a favore della distribuzione Binomiale Negativa e di una mistura di Binomiali Negative. La stessa cosa è stata fatta per la distribuzione di severity, per il cui "corpo" sono state provate la distribuzione esponenziale, lognormale e Weibull, mentre la coda è stata modellata attraverso la GPD applicando il metodo POT descritto in precedenza. Per fare ciò abbiamo individuato una soglia u, per il settore finanziario, pari a 200.000 records persi attraverso la visualizzazione del mean excess plot dei dati.

I parametri delle distribuzioni sono stati stimati attraverso il metodo della massima verosimiglianza. I risultati dei test condotti e dell'analisi grafica ci portano ad affermare che il modello più adatto per caratterizzare la severity dei nostri dati è quello costituito da un "corpo" lognormale con una coda GPD, dando vita dunque ad una spliced distribution; per quanto riguarda la frequenza, sia la Binomiale Negativa che la mistura di Binomiali Negative a cui si è accennato in precedenza, e che va a descrivere la doppia modalità degli eventi inclusi nel dataset, forniscono un fit adeguato.

Il passo conclusivo per la costruzione del modello consiste nell'aggregazione delle distribuzioni di frequenza e di severity stimate fino a questo punto, in modo da ottenere la distribuzione aggregata di probabilità delle perdite su un orizzonte temporale di un anno, Z. Si tratta di un'operazione che, in generale, può essere svolta in diversi modi più o meno efficienti; infatti, quando non è possibile ricorrere a formule analitiche ben definite, come in questo caso, si può fare affidamento su algoritmi numerici per ottenere un'approssimazione della

distribuzione aggregata. Tra questi, il metodo più diffuso è sicuramente il metodo Monte Carlo, che consiste nella simulazione di n scenari sulla base delle distribuzioni di frequenza e di severity, da "aggregare" successivamente in modo da ottenere una stima della nostra distribuzione d'interesse. In questa ricerca sono state stimate, simulando 10.000 scenari mediante l'algoritmo presentato nel testo, quattro diverse aggregate loss distributions, una per ognuno dei modelli severity/frequenza ritenuti adatti allo scopo; oltre ai modelli già accennati in precedenza, è stato testato anche un modello semplice costituito dalla sola distribuzione lognormale a caratterizzare l'intera distribuzione di severity, e un modello EVT-POT dove però è stato applicato un vincolo di continuità tra corpo lognormale e coda GPD in corrispondenza del valore soglia u. Attraverso l'analisi dei risultati della simulazione concludiamo che i modelli più adatti sono proprio quelli a cui si è accennato in precedenza, che forniscono le stime di rischio più ragionevoli, mentre il modello semplice e quello con vincolo di continuità sono da scartare.

Attraverso il modello si è dunque stimato un VaR al livello di significatività del 99,9% pari a circa 220.343.016 records per il modello EVT con la Binomiale Negativa singola; pari a 248.193.137 records per il modello che considera la mistura di Binomiali Negative come distribuzione di frequenza. I valori di rischio espressi non sono immediatamente convertibili in unità di misura monetaria in quanto stimare un costo medio per record è un compito alquanto arduo; attraverso l'applicazione del modello log-log stimato da Jacobs (2014) risaliamo ad un VaR 99,9% che oscilla dai 4,6 miliardi di dollari per il primo modello ai 5,034 miliardi per il secondo.

L'analisi è stata poi ripetuta anche sui sotto-dataset inerenti al settore medico e al settore "business" generico; le impressioni sulla diversa qualità dei modelli ipotizzati e stimati sono state quindi confermate anche da questa ulteriore applicazione. Si evidenzia come il settore medico, sebbene sia il più colpito in termini assoluti tra i settori compresi nel dataset e anche il meno "protetto", presenti un grado di rischiosità sensibilmente più basso se comparato con gli altri due settori considerati; al contrario, il settore business possiede il grado di rischiosità più alto, nonostante risulti anche quello meglio protetto. Il settore finanziario si va a porre esattamente a metà strada tra questi due settori, in termini di rischiosità.

I risultati ottenuti attraverso l'analisi svolta in questo lavoro sono da intendere come puramente illustrativi della dimensione e dell'importanza, anche in prospettiva futura, del fenomeno del cyber risk. È stata analizzata infatti solo una delle varie configurazioni in cui questo tipo di rischio può materializzarsi, e cioè i data breach, eventi di perdita accidentale o di

furto di dati sensibili, che secondo Eling and Wirfs (2018) costituiscono solo il 25% del totale dei cyber events.

Un'analisi illustrativa di questo tipo, per sua natura, non può essere esente da limitazioni teoriche e pratiche; si è già diffusamente parlato, ad esempio, del problema relativo alla mancanza, almeno su base pubblicamente disponibile, di dati consistenti su cui poter effettuare analisi statisticamente robuste. Il dataset utilizzato in questo lavoro infatti è costituito solo da eventi pubblicamente riportati dai media e quindi per sua natura incompleto. Un'altra importante limitazione riguarda poi l'eterogeneità degli eventi inclusi nel nostro dataset, che riguardano organizzazioni finanziarie di diversa natura, grandezza (sia in termini di espansione geografica che di dimensione economica) e importanza; per questo motivo, i risultati in termini di rischio ottenuti vanno intesi, come già sottolineato, come complessivi per l'intero comparto e non a livello di singole organizzazioni. Infine, è opportuno sottolineare anche come si sia scelto di aggregare gli eventi "accidentali" e quelli "intenzionali" analizzandoli insieme come parte di un unico profilo di rischio.

Pur ferme queste limitazioni, comunque, il lavoro presentato fa uso di una metodologia ben definita, di semplice comprensione, replicabile e adattabile, sperimentata con frequenza crescente negli anni dalla letteratura correlata e si ha dunque motivo di ritenere che con input di migliore qualità e ulteriore raffinamento, il modello creato attraverso tale metodologia possa in qualche modo rappresentare una valida alternativa.

Attraverso l'analisi svolta si è sottolineata l'importanza dell'applicazione degli strumenti della Extreme Value Theory per rappresentare adeguatamente un fenomeno in cui gli eventi di impatto estremo si verificano con una frequenza e una probabilità più alta di quanto i modelli probabilistici standard riuscirebbero a catturare; si è testato l'utilizzo di una distribuzione di frequenza "mista", che riuscisse a descrivere adeguatamente la frequenza delle due "modalità di eventi" presenti nel dataset, e si è infine accertata la significatività del modello anche per settori diversi da quello finanziario.